VOL. 15 No. 5, pp. 133-160 NOVEMBER 1953

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

MODERN CASTING TECHNIQUES

by H. J. MEERKAMP van EMBDEN. 621.74.045

Casting by the lost-wax process, an ancient craft practised with great skill already thousands of years ago by the Sumerians and the Egyptians, has retained its importance throughout the ages up to modern times. Many of the beautiful art bronzes of the Renaissance were cast by this method, and the same technique forms the basis of a modern industrial process for the mass production of metal articles, which has been adopted during the last decade.

The casting of molten metals to form articles of general utility is a very ancient craft that goes far back into the Bronze Age. If we examine the history of metallurgy, we find that the first metals to which primitive man had access were found in a native state (gold, silver, copper.) At a very early period, however, copper and bronze were separated from their oxidic ores in a primitive fashion, and the metals thus obtained were shaped by means of casting, sometimes followed by forging and other processes, into weapons and objects for religious and domestic use. Cast copper and bronze articles are known to have been already made in various regions in the third or fourth millennium before Christ.

The metallurgy of iron underwent an entirely different development. Native iron is very rare (meteorites), and the melting-point of iron was too high for the craftsmen of those days, so that extraction was restricted to reduction of the ore to metal in a solid or dough-like state. This metal, impregnated with slags, was hammered into bars by repeated forging, which also had the effect of expelling the slag. Only when the development of the furnaces was in an advanced state did it become possible to obtain liquid iron. The oldest known cast-iron comes from China and dates from about 500 B.C. In Europe the great technical development of cast-iron begins towards the end of the Middle Ages. Cast-steel, owing to its low carbon content, has a melting point lying 350 °C higher (above 1500 °C), and could be made only about the beginning of the 19th century, while the modern alloyed stainless and heat-resistant steels have only in the last few decades begun their remarkable development, as is the case with aluminium, magnesium and their alloys.

The number of alloys employed has greatly increased, especially in this century, and the same can be said with regard to the number of melting and casting methods. The crucible furnace, in which the material is melted in a crucible, the reverbatory furnace in which burning gases pass over the material, the blast furnace in which ore and coal are introduced together at the top while the air required for the blasting is blown in and the metal and slag drawn off from below: these are the furnaces which have been developed in the course of the centuries to the perfected equipment of to-day. Quite new are the electric furnaces. The introduction of the electric-arc furnace was greatly stimulated by the first World War. The high-frequency furnace, which made its début in industry during the 1930's plays only a very limited part in production as yet, being used mainly in the production of high-alloy steel. The low-frequency induction furnace, on the other hand, has won for itself a wide field of employment in the melting of

brass and bronze. Aluminium and magnesium alloys are very often melted in electric radiation furnaces.

Casting in sand-moulds

Now for the actual casting methods. Originally, a pattern was pressed in sand or loam and the hollow obtained was filled up with molten metal. At an early date, however, castings were made in closed moulds: two boxes filled with sand in which a pattern (usually a wooden one) was laid, so that a hollow was formed, half in the lower box and half in the upper box (*fig. 1*). The top box was then lifted off, the pattern removed and the top box replaced on the lower one. In the top box a channel ("sprue", "runner" or "gate") was made through which the metal was poured. After solidification of the metal, the mould is "shaken out", and gates and runners are cut off. During casting, the air in the mould and the water-vapour and gases liberated from the mould material by the hot metal, escape through the porous mass.

In this way, intricate castings could be produced. The technique has been more and more perfected in the course of the centuries, until we come to the modern sand or loam casting by which, with the aid of all kinds of contrivances, complicated parts can be cast, that form the basis of our tools and machines. No great dimensional accuracy can be expected, however, for two reasons. The first is that, after removing the pattern, it is not possible to replace the halves of the mould so accurately that there is absolutely no alteration in their mutual position. The second reason is that the pattern has to be taken out of the sand, which entails some knocking or vibrating to loosen it, with the result that the dimensions of the mould are slightly different from those of the pattern itself. A restriction on the shape of the moulding is that the pattern must be "detachable", that is, it may have no protruding parts that would prevent its removal from the sand. Yet it is astonishing what complicated castings can be made by the subsequent insertion of loose cores of sand fixed in the hollow of the mould. The mould for the cylinder block of a motor-car, for instance, is built up of a large number of sand parts (sometimes more than fifty). In this case various kinds of binding agents are added to the "sand" (clay, oil, cement and numbers of others) in order to meet all requirements as to strength, porosity (for the escape of gases etc., see above) and the quality of the metal surface.

Fig. 1. Upper and lower boxes of a sand mould. The wooden pattern can be seen in the foreground. In the upper box the "gate" or "sprue" can be seen through which the metal is poured. (This photo, taken in the Reineveld Engineering Works at Delft, was supplied by the Casting Research Centre, National Council for Industrial Research, Netherlands.)

Permanent moulds

Interesting as the subject may be, we shall not go further into "sand-casting" but turn our attention to other methods. One of the oldest employs permanent moulds that can be used more than once. Even primitive peoples made use of such moulds hewn in stone and this method developed into metal moulds that are still employed on a large scale. They are, for example, used for the casting of ingots. These are afterwards worked up into sheets, bars, strips and wire by hammering, forging, rolling etc., thus forming the raw material of our entire metal working industry. Quite 90% of the metals at present consumed by industry is worked up in this way and feeds the metal working machines, and only a small part of the total supply of metal is cast directly to shape.

Permanent moulds are also used for the direct manufacture of utility articles. We speak of "gravity die casting" if the metal is poured in with a ladle, and of "pressure die casting" if the metal is forced into the die under (usually high) pressure (*fig. 2*). Gravity die casting is an old process; pressure die casting is an invention of this century of mass production[1]). This method is especially employed for the mass production of small parts made of low melting-point material. An example is the zinc alloy, from which all kinds of automobile parts are made, from the intricate carburettor to the radiator grill and the doorhandles. In addition, a great deal of pressure die casting is performed in aluminium and brass. Metals with higher melting-points cannot in practice be cast in this way, as the die cannot stand the sudden change in temperature during the injection of the metal, and is soon damaged.

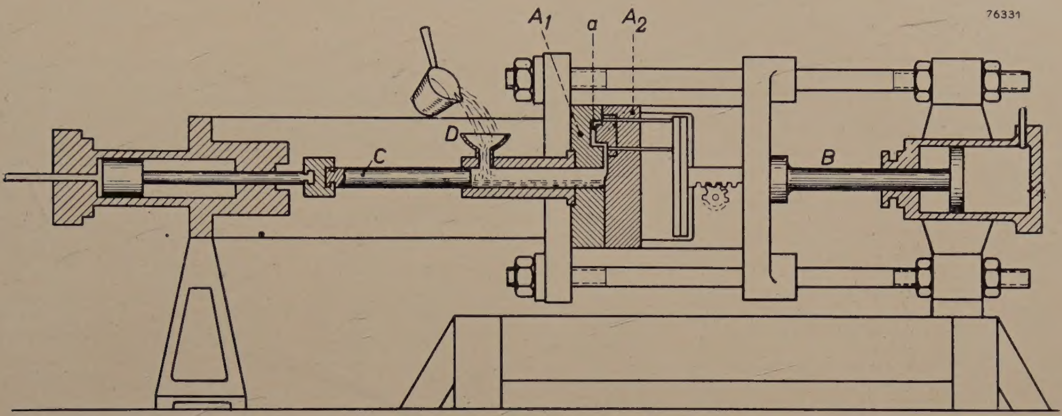


Fig. 2. Machine for pressure die casting. A_1 and A_2 are the two parts of the steel die. The moving part is pressed against the fixed part by the hydraulically moved piston B, while the piston C forces the liquid metal, poured into funnel D, into the cavity a of the die. (From "Designing for Alcoa die castings", Alum. Co. of Amer. Pittsburgh 1948.)

Besides these methods, two others have recently come to the fore: the first is the so called "Croning process" (C-process or shell-mould casting) and the second is the lost-wax process, also called "precision casting" and "investment casting" (see later).

Croning process

The Croning process[2]) is new and originates from Germany where it was invented during the last World War. It it based on the following principle (cf. *fig. 3*). A metal pattern is divided into two

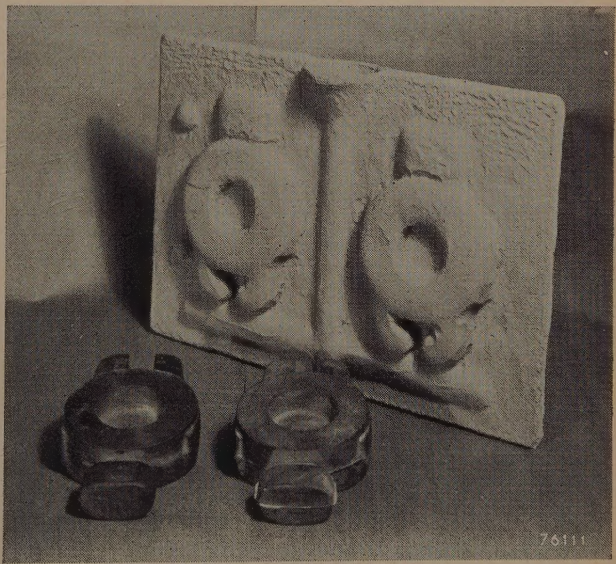


Fig. 3. Half of a double mould for the Croning process. The two halves of the mould are made by dropping a mixture of fine sand and powdered plastic on a heated pattern. In the foreground are pieces cast from such a mould. (From K. Rose, Materials and Methods, Jan. 1953.)

[1]) See for example. "Die casting for engineers", The New Jersey Zinc Co., 1946.

[2]) Field Information Agency (Technical), Final Report 1168, "C"-process of making moulds and cores for foundry use.

halves and each half is mounted on a flat plate. The plate is heated to about 250 °C and then placed face downwards on a box containing a small quantity of moulding powder. This powder is a dry mixture of fine sand and powdered resinous plastic. The box with the plate is then turned over, the powder falls on the hot mould, the plastic melts and causes the grains of sand to cohere. After a few minutes the box is turned up again and the pattern plate, on which a "shell" of coherent sand about $^1/_2$ cm thick is formed, is taken off the box and heated. The resin-sand shell is "cured" by the heating and after some time can be separated from the pattern. Two of these shells are fitted together and in this way form the mould into which the metal can be poured. During casting the mould becomes so hot that after the solidification of the metal the resin will for the greater part be burnt out, so that the remains of the mould can easily be removed from metal. The fine sand gives a clear impression with good detail. The process has proved itself to be very suitable for the mass production of small articles. The United States in particular are attempting to perfect the technique [3]). Many industries are developing the process to extend its use to the most varied products of cast-iron and other metals. The croning process is even being developed for certain high-grade steels, e.g., for gas turbine parts.

Lost-wax casting

The last process that we shall examine here is lost-wax casting. This will be dealt with rather more thoroughly, in view of the investigations which have been made in the Philips laboratories into the possibilities of its utilisation. The process

76001

Fig. 4. Ancient Egyptian bronze (from the collection of the late Dr. A. F. Philips, at Eindhoven). The bronze is about 15 cm high. These bronzes were cast by the lost-wax method. In excavated workshops, the remains of wax patterns have been found.

[3]) R. W. Tindula, Foundry, Cleveland, **80**, 201-210, 1952 (No. 7).

itself is very old. The Sumerians and the Egyptians practised this method already three thousand years before Christ [4]). Decorative objects were first fashioned in wax. These were then covered with wet clay in which a hole ("sprue") had been left. The lump of clay was then dried and afterwards heated, causing the wax to melt away. The molten metal was poured into the cavity thus obtained. Many bronze images and objects that accompanied the Pharaohs on their journey to the realm of the dead were made in this way [4]), and exhibit a great sense of art and great skill (*fig. 4*).

The great value of this process for the plastic arts is obvious if we compare it with sculpture in hewn stone. The free modelling in wax, the ease with which smooth surfaces and fine details can be reproduced, the mechanical strength of work cast in metal and finally the possibility of forming the wax pattern (which is lost during the process — hence the name) by means of a mould, whereby the process is potentially repetitive: all this induced artists in many cases to prefer metal casting to working in marble or stone.

The technique has therefore remained in use throughout the centuries for the making of bronzes. It reached a culmination in the Italian Renaissance. Vasari, in the prologue to his well-known "Lives of Artists" (1550), gives a detailed description of the technique [5]). Benvenuto Cellini (1500-1571) the goldsmith-sculptor, renowned for the splendid bronzes, some of them of great dimensions, which he created by this means (*fig. 5*), has written an extensive record of the methods he employed [5]).

During the last 50 years, however, except for sculptors in bronze, it was only the dentists who used wax impressions, to serve as a pattern for the making of the gold castings for false teeth! Not until about 1930 did the process find any other employment (viz., in the jewellery industry) but about 1940, when the gas turbine made its début in aircraft, the process was introduced on a large scale. The blades of gas turbines (*fig. 6*) have to come up to the highest mechanical and thermal requirements, and the newly developed materials of which they were made were difficult to machine.

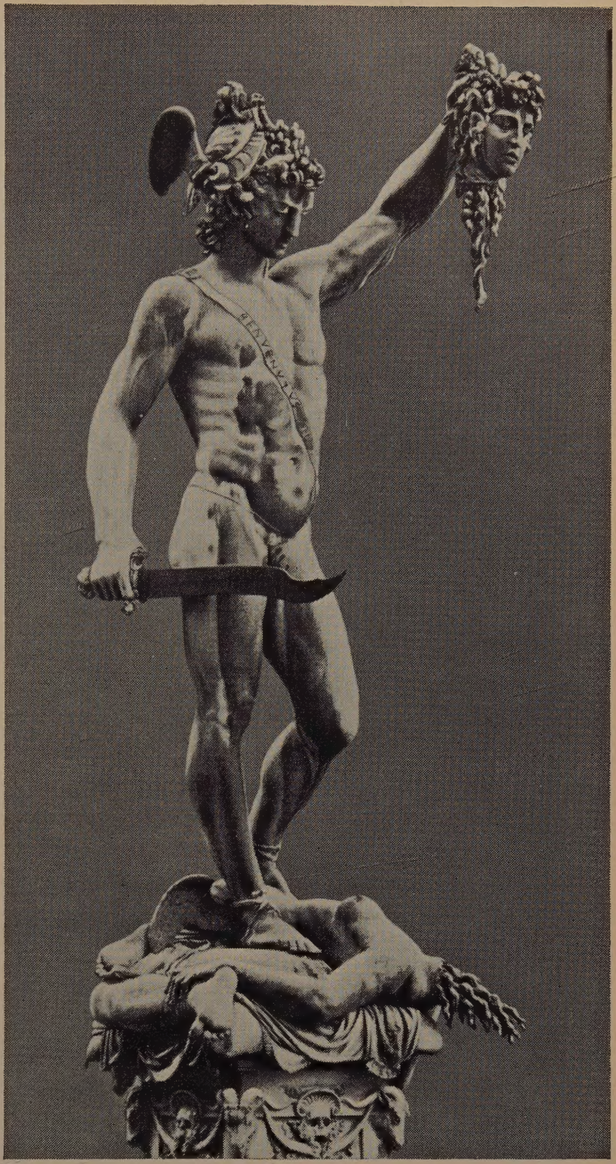


Fig. 5. Benvenuto Cellini's bronze Perseus with Medusa head, in the Loggia dei Lanzi at Florence (completed 1545). The statue displays a great fineness of detail, made possible by the lost-wax casting method. (Block kindly supplied by A. Oosthoek Publishing Co., Utrecht.)

Lost-wax casting was the solution of this problem. At the end of the war there was an American factory with a monthly production of two million blades! It is small wonder that after its successful application to the gas turbine, the process was tried for other products as well, and is now employed in many countries, though still on a limited scale [6]). Some years ago Philips set up a research group to

[4]) See for example: H. Garland and C.O. Bannister, Ancient Egyptian Metallurgy, Griffin, London 1927.
G. Roeder, Die Herstellung von Wachsmodellen zu ägyptischen Bronzefiguren, Z. ägypt. Sprache und Alt. **69**, 45-67, 1933.
V. Gordon Childe, The technique of prehistoric metalwork, Antiquity **22**, 29-33, 1948.

[5]) G. Vasari, Le vite de' più eccellenti architetti, pittori e scultori, Florence 1550.
B. Cellini, Trattati dell' oreficeria e della scultura, Panizzi and Peri, Florence 1568.

[6]) L. G. Daniels, Great savings in replacing complex assemblies by a single investment casting, Metal Progress **56**, 490-491, 1949.
R. L. Wood and D. V. Ludwig, Precision investment castings replace parts produced by other methods, Materials and Methods **32**, 49-53, 1950.
R. Miller, Investment casting, money saving mass producer, Steel, 14 Jan. 1952.

study this casting technique and put it into practice. The process has now been in use for some time on a factory scale. The photos *fig. 7a* and *b* give an impression of some of the machine parts made here.

We shall now discuss further the present industrial form of this process, in connection with the manufacture of machine parts, particularly from a mass production point of view.

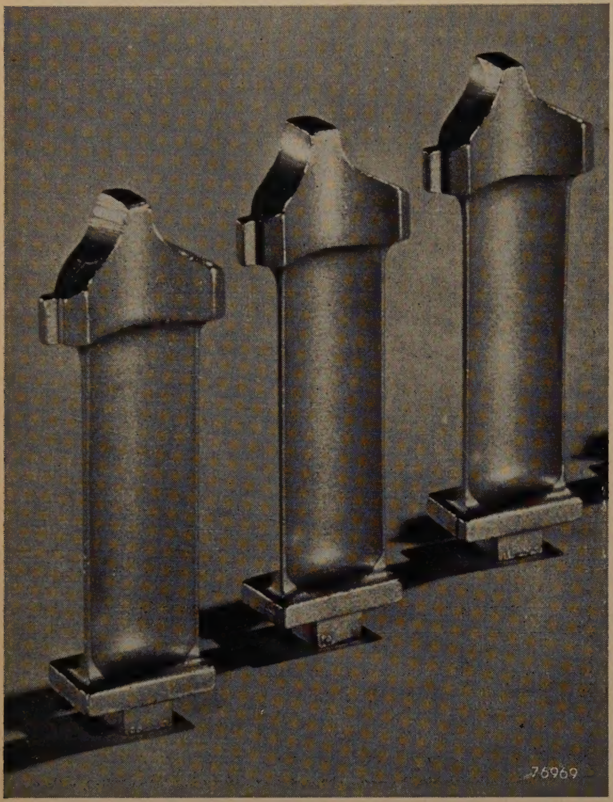


Fig. 6. Blades for gas turbines cast from a heat-resistant alloy by the lost-wax casting process. The blades are about 4 cm long. (By the courtesy of Haynes Stellite Co., Kokomo, Ind., U.S.A.)

The procedure is as follows: a wax pattern is first made of each article to be cast. These patterns are then mounted on a "wax-tree", the trunk of which will serve later on to pour in the metal. The "tree" is mounted vertically on a wax covered plate, and a box or cylinder without ends is placed on the plate and melted to it. The lidless box formed in this way is filled with a "slurry", a ceramic mass that solidifies after some time. (This operation is called investment: hence the term "investment casting".) The box and its contents are then heated, the bottom plate loosens and is removed, and the wax melts and runs out of the mould, which is then dried and baked. The mould is now ready, and the cavity is filled with molten metal. After cooling, the mould is shaken out, the castings are cut off the "tree", cleaned and, if necessary, machined.

After this introductory description we shall examine the various steps in the process more closely and then consider what the advantages and limitations of the process are.

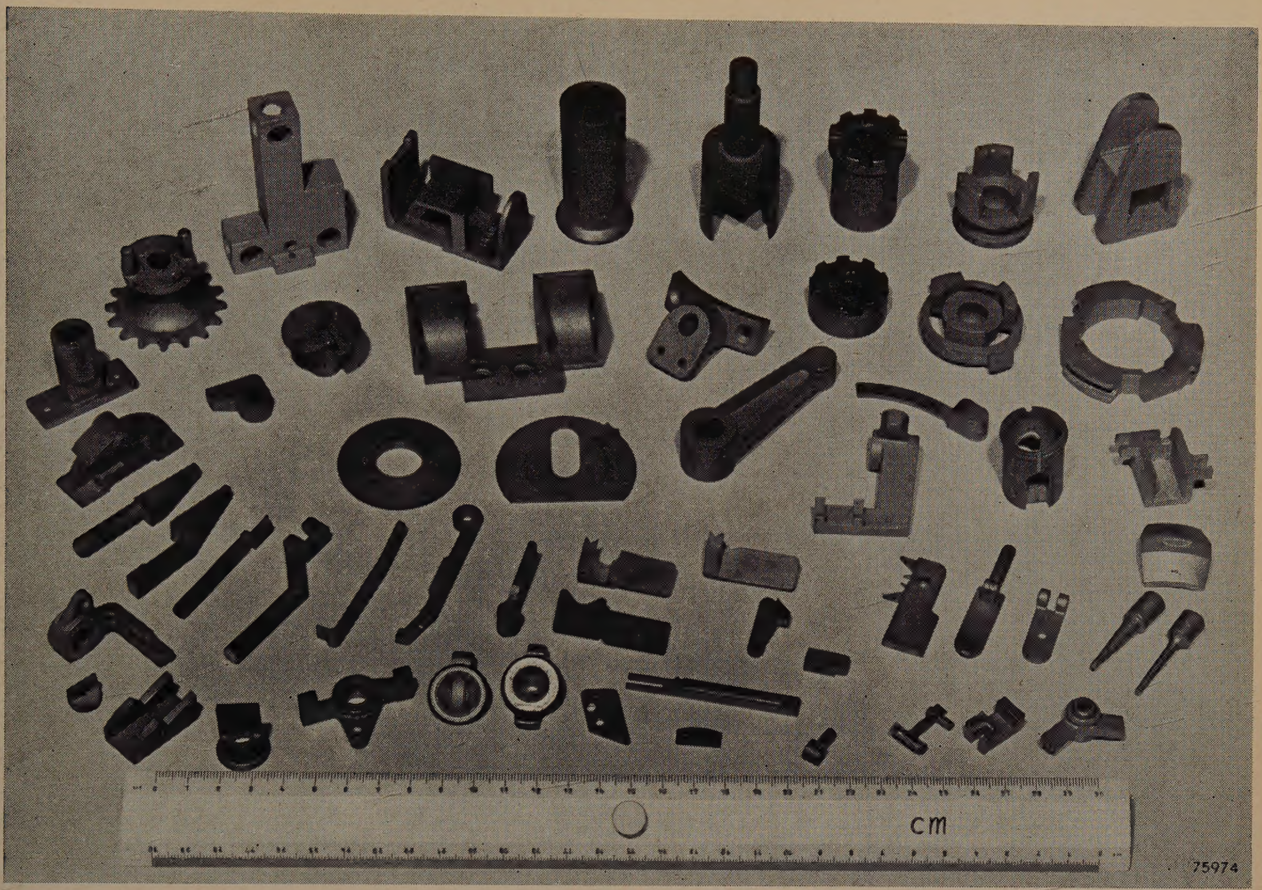
Making the wax pattern

For making the wax pattern we need a metal die. This is filled with liquid wax, and opened, after the wax has set, to allow the pattern to be removed. The wax pattern can also be composed of various parts made by means of separate dies and then simply joined together with a not-too-hot soldering iron. The stipulation that the pattern must be detachable (from the die) now applies only to each part, and by a suitable subdivision of the wax die, this last restriction on the shape of the object can be eliminated. (Figs. 8 and 9 serve to illustrate this.)

A die, consisting usually of two halves (for intricate wax patterns, it may consist of several parts), can be made of steel or brass by the normal manufacturing methods. Another method is to start from a so-called master pattern, a hand-made metal pattern of the object to be cast. This is embedded in a metal of low melting-point (preferably an alloy of bismuth and tin, which undergoes practically no change in dimensions when solidifying). This is carried out in two parts, first the lower half and then the top half, so that the die can be opened (*fig. 8*). The two halves are then pegged and an inlet is drilled for the wax. This method is especially attractive for irregular patterns such as gas turbine blades, but the bismuth-tin alloy is rather brittle, so that damage often occurs and the sharp edges of the die soon crumble. For this reason the dies are often made according to the former method, that is of brass or steel (*fig. 9*). The first is easily machined but rather soft, while steel has the disadvantage of rusting and is more difficult to machine. The better the finish of the die the better the wax pattern will be and hence the better the subsequent casting. It is therefore better to pay greater attention to the making of the dies if quality products are required.

One of the difficulties that does not arise in casting art objects but is at once obvious in the case of industrial castings, lies in the fact that the ultimate product does not possess the exact dimensions of the wax die. There is a certain shrinkage. The wax pattern must therefore be about 1 to 2% larger than the finished product. Unfortunately the shrinkage is not equal in all directions and must therefore often be determined empirically, which entails corrections to the die or to the master pattern. If the dimensions have to be very accurate, as in the case of turbine

a



b

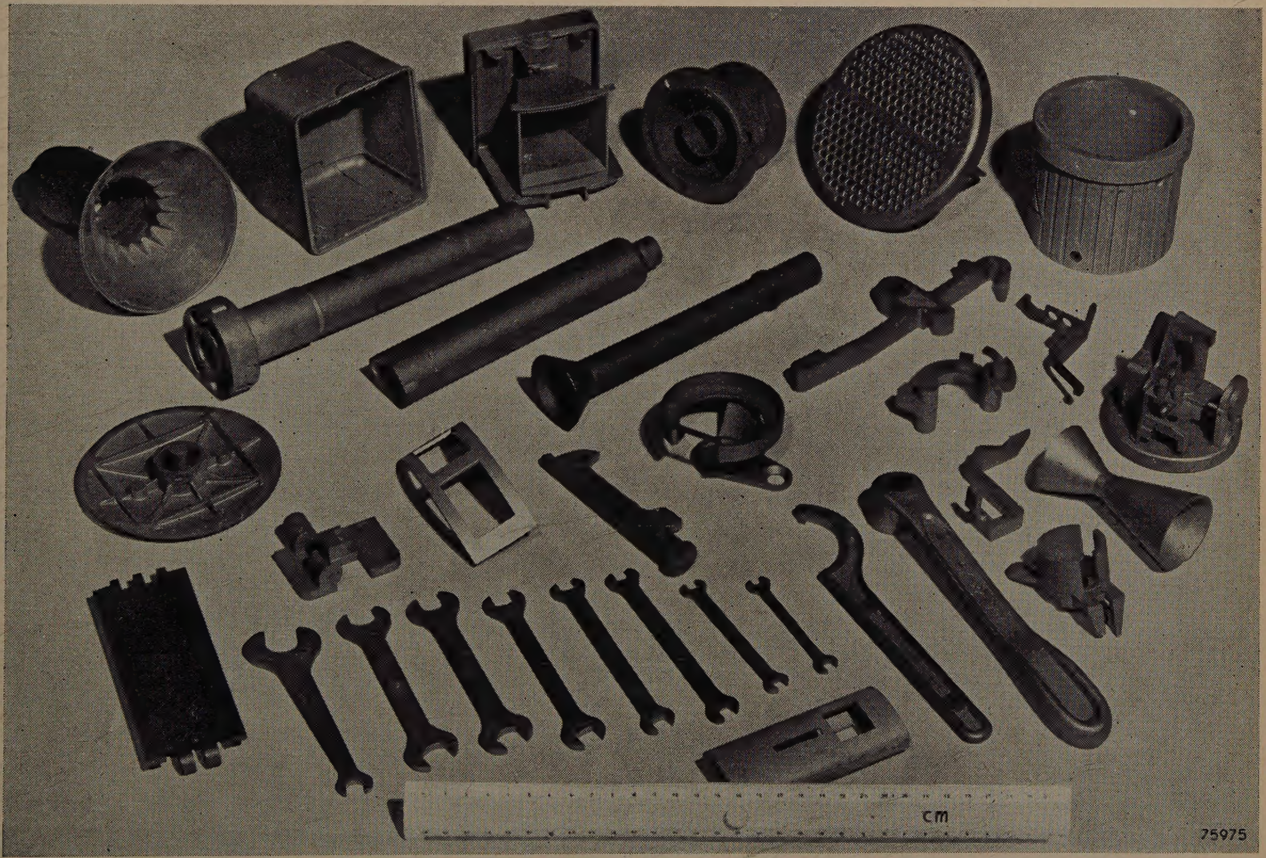


Fig. 7. *a*) and *b*) A selection from the series of tools and parts cast in the precision casting department of the Philips laboratories at Eindhoven. None of the pieces shown has been machined. Note the wide variety of shapes and the smooth surfaces. These articles are cast from a number of different alloys.

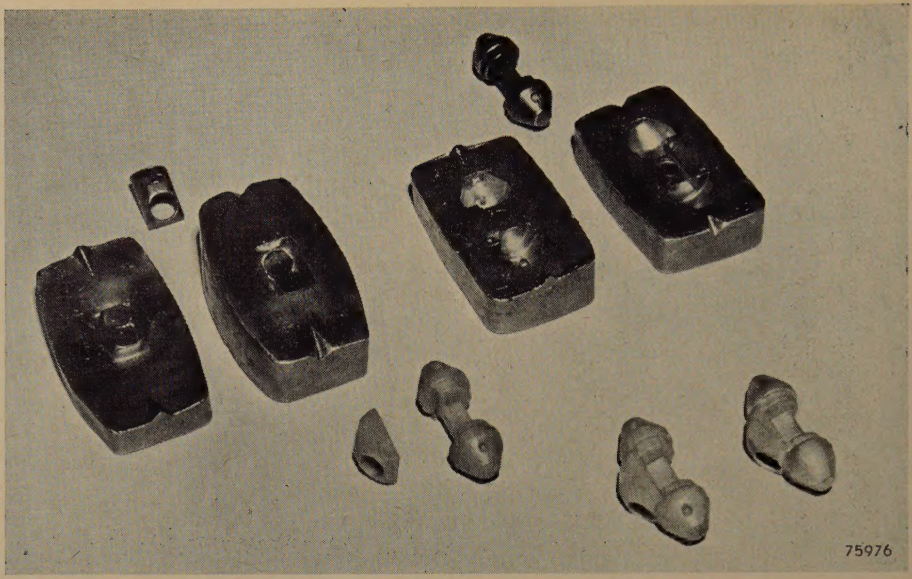


Fig. 8. Wax die of bismuth-tin alloy. Above: the hand-made master pattern, composed of two individual patterns. Below: bottom and top half of each of the two dies made by embedding the two master patterns in the molten bismuth-tin alloy. Further below are the wax patterns made with these two dies, and a complete wax pattern obtained by "soldering" two of these patterns together. At the bottom right is an aluminium casting made with such a wax pattern.

blades, the following procedure is followed: a master pattern is made, from this a wax die and from this latter a trial series of wax patterns which are finished normally into blades. From the results obtained, it can be decided whether the master pattern can be corrected or whether a new pattern must be made. Then a second and, if necessary, a third trial is made, until all requirements are fulfilled and production can begin.

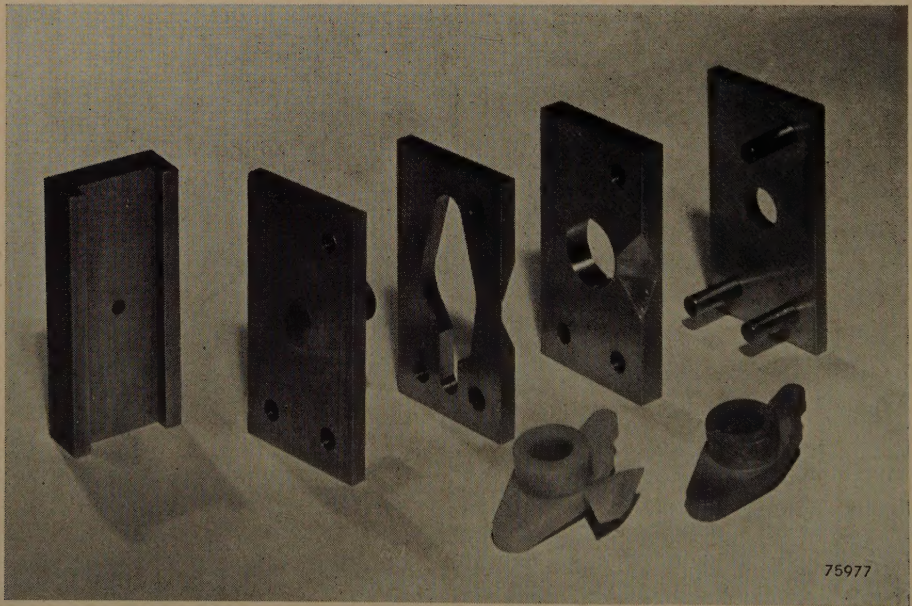


Fig. 9. Wax die of several parts made of brass ("exploded view"). In the foreground is a wax pattern (with gate) made from the wax die, and a casting (chromium steel).

The press which extrudes the wax from a reservoir into the dies (*fig.10*) usually has electric heating elements, automatically controlled to ensure uniform temperature. It is often operated by a compressed air supply at a pressure of 4 to 5 atmospheres.

The wax used is usually a mixture of various species of wax and paraffins. The exact composition is chosen with some care so as to obtain a reproducible product. The wax, when solidifying and cooling, must not undergo much shrinkage and must be hard when cold, otherwise the patterns would sag

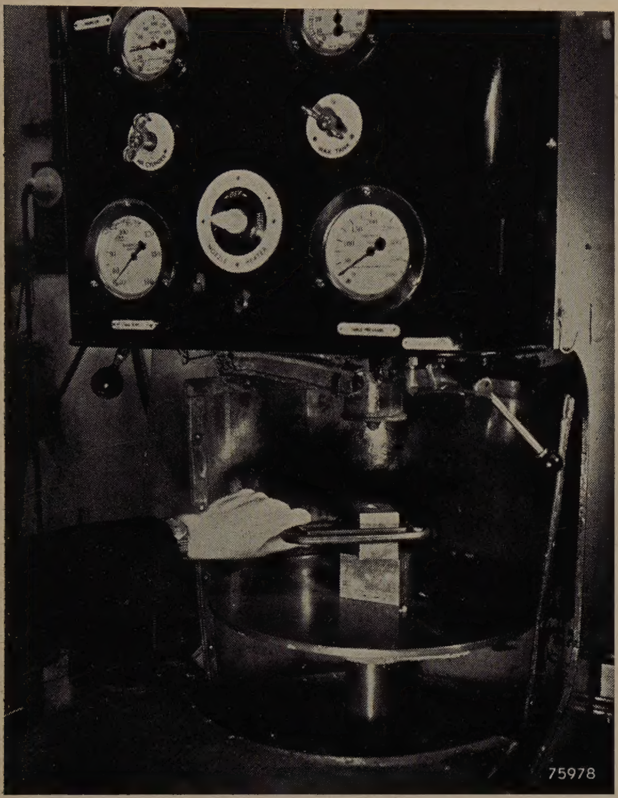


Fig. 10. Wax press. The wax die, its two halves clamped firmly together, is placed on the hydraulically moved table; the die is about to be raised and pressed with its opening against the head of the nozzle, after which it will be filled with wax from a heated reservoir under a pressure of several atmospheres.

even before casting. Instead of wax other materials are also used. Certain plastics (chiefly polystyrene) are often used if a large number of castings is to be made [7]). This material is much stronger, so that the patterns made of it do not easily break, but it has to be pressed in a die under very high pressure. Instead of the simple wax dies that withstand a pressure of five to fifty atmospheres, dies of hardened steel and pressure die machines capable of handling some thousands of atmospheres have to be used. After investing, the plastic patterns are removed from the mould by burning them out.

Frozen mercury is sometimes used instead of wax or plastic. This requires temperatures of — 40 °C (below the melting-point of mercury), the technical problems concerning which have apparently been solved. The advantage of this method, according to the inventor, is a very fine sharp impression of the (steel) die, because the mercury remains liquid when poured in and fills the die well, and has a low shrinkage when solidifying [8]).

Making the mould

The "treeing" of the wax patterns is done fairly simply with the aid of a soldering iron (*fig. 11*). In many cases the wax die can be so constructed that the gate of the mould is already formed, and for small, simple articles, the wax die can be provided with a number of cavities, so that with one injection of wax a number of wax patterns can be made. Generally the same requirements obtain here as in the moulds for sand castings: the metal must flow into the mould as uniformly as possible without sudden strictures or changes of direction that might hinder the steady filling of the mould with metal.

After the tree has been stuck on to the waxed plate and enclosed by the lidless box, the latter is filled (*fig. 12*) with the moulding mass. This is a slurry poured round the wax patterns, which after some time solidifies or "petrifies". When the mould is petrified, the loose bottom is removed and the mould warmed. The purpose of this is to dry the mould and to melt out the wax. A relatively low temperature is sufficient for this. Later on, when the greater part of the moisture has evaporated, the temperature is gradually raised up to 1000-1100 °C to drive out the remaining moisture and to burn out the remnants of the wax in the mould.

For the preparation of the moulding mass or slurry, various materials have been used. Plaster of Paris can be employed, sometimes mixed with quartz powder or other ceramic materials. Often aqueous suspensions of quartz powder and ground refractories, with silicic acid as a binding agent, are used. This transforms the liquid after some hours into a gelatinous mass, binding the suspended particles to a fairly firm mass. Ethyl silicate is usually the basic material, which on decomposition is converted into silicic acid and ethyl alcohol. Sometimes the cheaper alkali silicates are used. Use is also made of the Sorel reaction, based on the action of magnesium oxide (practically insoluble in water) on phosphoric acid or acid phos-

[7]) W. H. Sulzer, Das Präzisionsgussverfahren, Neue Giesserei **37**, 557-565, 1950

[8]) W. I. Neimeyer, Precision casting with frozen mercury patterns, Iron Age **163**, 94-97, 1949.
H. Chane and L. T. Schakenbach, New precision casting process, Materials and Methods **29**, 52-56, 1949.

phates. A slow reaction takes place, and in the course of some hours, basic magnesium phosphates are formed, which take up large amounts of water, so that the slurry solidifies after several hours.

With these materials a rather thin slurry is obtained, which entirely encloses the wax patterns and which, after the escape of the enclosed air bubbles, solidifies. The water disappears in the drying process and now it is important to keep shrinkage and distortion to a minimum. For this reason the box with the still-liquid mass is usually placed first of all in a vacuum (*fig. 13*), thus enlarging the enclosed air bubbles so that they rise more easily and rapidly. It is then vibrated on a vibration table (fig. 12). This causes the solid particles in the mass to sink, and a dense packing of the grains is obtained before the liquid becomes thick and gelatinous. When the mould is subsequently dried and heated (baked), the coarse grains touch each other and the mould remains practically free from shrinkage. There are of course cavities in the spaces between the grains, so that the mould remains porous when hard.

Before the moulding mass is poured into the boxes, a thin, not very porous coating is applied to the wax patterns. When the metal is poured in, it would otherwise penetrate the fairly large pores of the baked moulding material, which is to be avoided. If the whole mass were made of so fine grain that no metal can penetrate, the porosity would be too low and the enclosed air and gases would be unable to escape in time. For this reason the wax patterns are dipped or sprayed with a thin coating of a very fine fireclay, dense enough to prevent the metal from penetrating and yet sufficiently permeable to allow the gases to escape rapidly. When this is done the mould is filled, vibrated and allowed to solidify and then, as already mentioned, the wax can be removed and the mould baked.

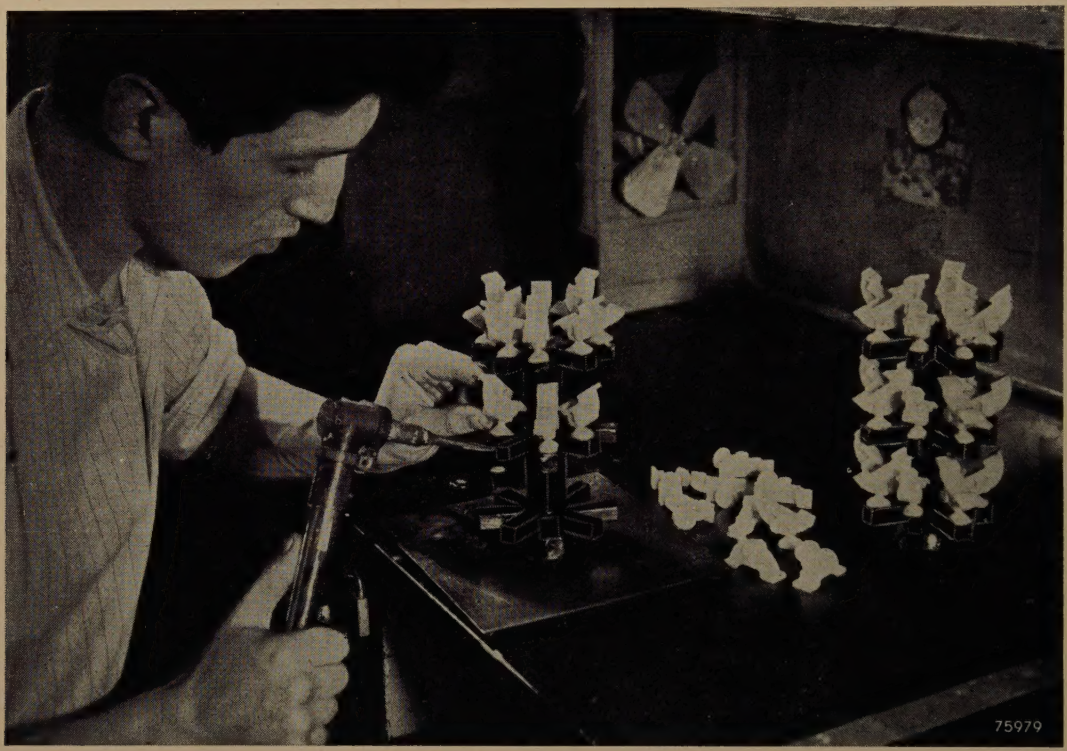


Fig. 11. "Treeing". The "sprues" through which the metal will flow have been formed beforehand (skeleton of dark wax). A great many wax patterns are "soldered" on to this. On the right is a completed "tree".

Filling the mould with metal

After baking, the mould can be filled with metal. It is desirable to keep the mould as warm as possible, so that the inflowing metal cannot solidify against the wall of the mould before it is full. The result is that castings can be made with far thinner walls than if the metal is cast in cold moulds. For castings of thick section, high mould temperatures have disadvantages, for the solidification

takes too long and a casting is obtained of coarse structure, which generally has poor mechanical qualities.

To fill the mould, the molten metal can simply be poured in, but it is better to speed up the process, again with an eye to possible premature local solidifying of the metal. Various methods can be applied for this purpose, e.g. applying pressure on the column of liquid metal, suction on the porous mould or centrifuging the mould. The method used is closely bound up with the method by which the metal is to be melted. For metals with a low melting-point, such as aluminium and magnesium, practically any type of furnace can be used, as the temperature of the melt does not exceed 700 °C. A large quantity of metal can be melted at once and the various moulds filled with a ladle, as is usual with gravity die casting. In the case of iron and steel, the melting temperature is so high that special furnaces have to be used. Generally two types are employed: the electric arc furnace and the high-frequency furnace.

Fig. 13. Before vibration the air bubbles are removed from the moulding mass by placing a vacuum bell over the box and connecting it for a time with a vacuum pump.

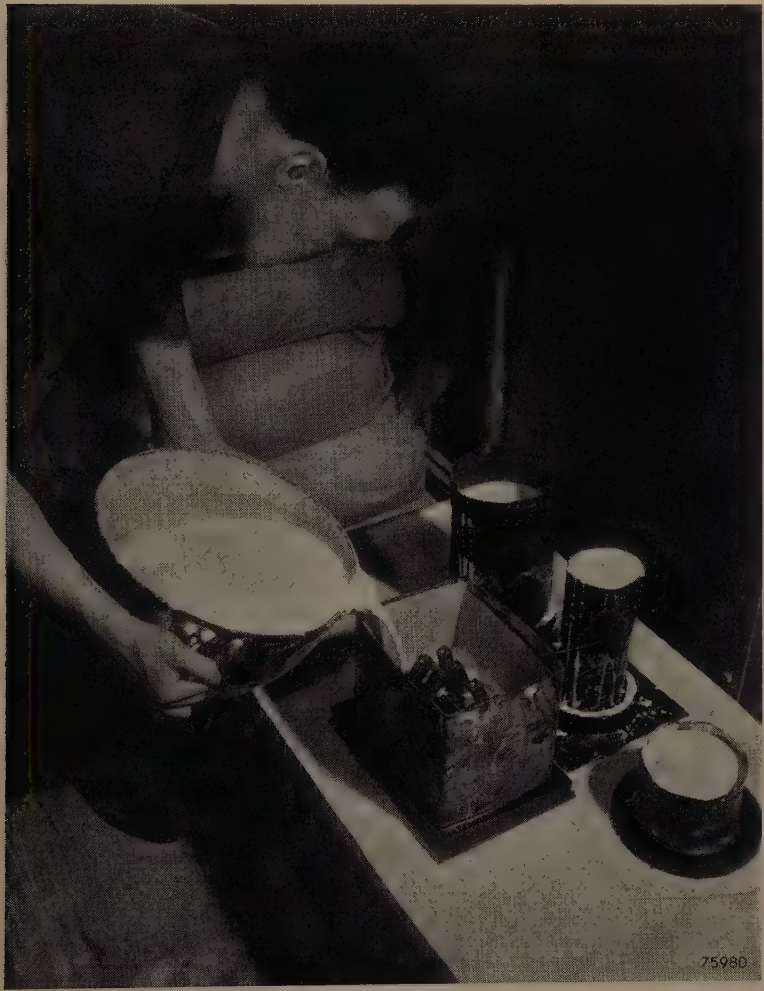


Fig. 12. The wax tree has been placed in a box and is being invested with the slurry. A number of boxes already filled, stand together on a vibrating table.

With the first type (*fig. 14*) the melting is performed in a refractory crucible enclosing a bottle-shaped melting space. The metal is fed through the neck. In the melting space, two graphite electrodes protrude, between which an electric arc is drawn. The radiation is reflected by the walls to the material to be melted, and the neck must be narrow to minimize the loss of radiation to the outside. One to three kg of metal can be melted in such a furnace, just enough for one mould as used in the present practice of lost-wax casting. This mould is then placed upside down on the furnace and clamped to it, after which the furnace and mould are inverted. The molten metal runs into the mould and can be forced in by connecting the furnace with a compressed air line.

The high-frequency furnace is slightly simpler in form. In this the crucible is surrounded by a water-cooled copper coil through which a high-frequency alternating current is conducted (*fig. 15* and *16*). The metal in the crucible is heated and fused by induction currents. The apparatus is on the whole more expensive than an electric arc furnace but the crucible can be cylindrical and can therefore be more easily filled with coarser pieces than the rods passing through the neck of the electric arc furnace. Scrap material and cut-off

"gates" from former castings can be fed into this furnace, and in this way save material. Moreover, the electric arc furnace is sometimes undesirable, for the melt is polluted by sparks from the carbon tips. With some materials, for instance stainless steel, this may be prohibitive, as even small quantities of carbon can have a very unfavourable influence on its resistance to corrosion.

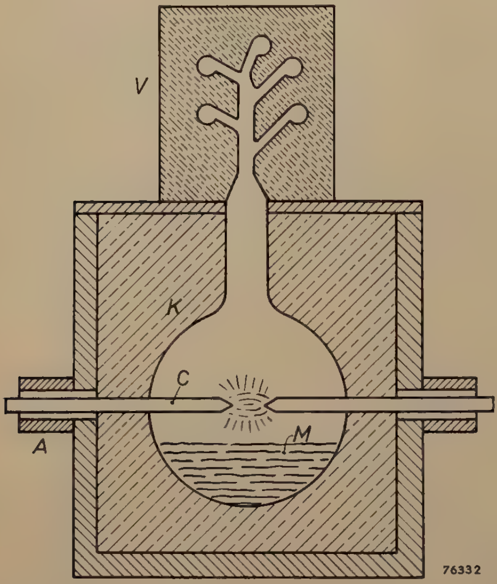


Fig. 14. Schematic section of an electric arc furnace. The metal *M* at the bottom of the bottle-shaped crucible *K* is heated to melting-point by the radiation of the arc between the two horizontal carbons *C*. The mould *V* is placed on the furnace and clamped to it. The mould is filled by inverting the furnace about the axis *A*.

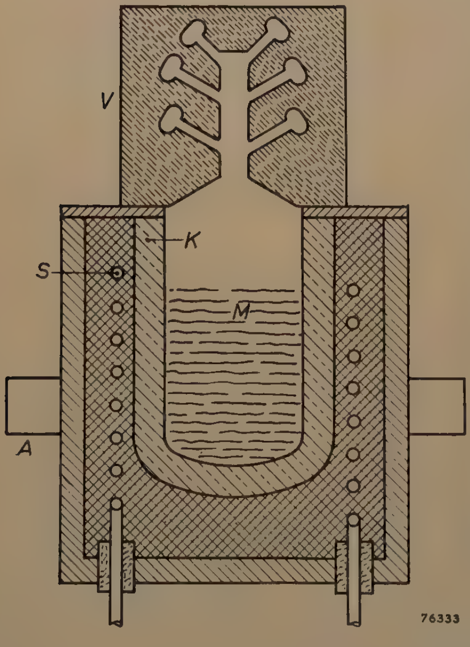


Fig. 15. Schematic section of a high-frequency furnace. The crucible *K* is here cylindrical. The load melts by the eddy currents induced in it by the high-frequency current sent through coil *S*. The wide mouth makes this furnace easy to feed.

After casting and cooling, the mould is shaken out, the "gates" are sawn or cut off and the casting cleaned and finished. Usually the finishing is done by sand or shot blasting and, if necessary, the metal may be heat-treated.

Advantages and limitations of the lost-wax process

The articles made in this way are not always ready for use. They often have to be machined in the usual way by turning, milling, drilling or polishing: the dimensional tolerances for modern machine parts are often much closer than can be attained by the lost-wax process. It is therefore a pity that the term "precision casting" has been invented for it, for "precision" is a very relative idea. Compared with normal sand casting, the dimensional fidelity and the quality of the surface are certainly much better, which may be considered great advantages over conventional casting methods. For small dimensions up to 50 mm (2″), tolerances are generally given of $\pm$ 0.2 or 0.1 mm ($\pm$ 0·008 or $\pm$ 0·004″), while some manufacturers guarantee $\pm$ 0.05 mm ($\pm$ 0·002″), but this is mostly effected by means of an expensive method of selection in which the samples not falling within the tolerances are rejected. For larger sizes, above 100 mm (4″), a tolerance of $\pm$ 0.3 mm to $\pm$ 0.5 mm (0·012″ to 0·02″) must generally be accepted. The limitations on the attainable precision are not surprising, having regard for the fact that the ultimate deviation is determined by the sum of a considerable number of factors. The wax pattern shrinks in the die, the dimensions of the cavity are affected by the expansion of the wax in melting and by the changes in dimensions during the drying and solidifying of the mould. The temperature of the mould when casting and the temperature of the cast metal, the shrinkage during solidification and cooling down to room temperature, sometimes obstructed by the resistance of the ceramic mould, all these co-operate in causing variations in dimensions. These can be reduced by working under the most constant conditions with regard to temperature and composition, but the tolerances just mentioned are, in practice, the utmost that can be expected.

In the mechanical workshop much smaller tolerances are implied by the term "precision work", viz., $\pm$ 0.005 mm, (0·0002″), or in special cases $\pm$ 0.003 mm (0·00012″); in this case, one is inclined

to reject even lost-wax castings with contempt, as the work of a blacksmith! But if the matter is examined more closely it will be noted that these high demands are rarely necessary for *all* the dimensions; a little machining will usually make the casting quite acceptable without much trouble. This is a matter of mutual arrangement between the designer, draughtsman and caster, are required with thin walls. Making such pieces from bar stock, from which large portions have to be removed by machining, is generally an expensive business, especially if materials have to be used that are difficult to machine. This is generally the case with wear-resistant and also with many corrosion-resistant alloys, and such materials are becoming more and more used in modern fabrication.

Fig. 16. High-frequency furnace, with casting mould attached, being tilted for casting. In the background is the high-frequency generator.

The most important advantage of lost-wax casting and of the modern casting methods in general lies, however, in our opinion, in the fact that now many kinds of alloys can be used which were hitherto out of the question. In general engineering the tendency is, rightly, to try to limit the number of alloys as much as possible. This will be readily understood, for if a certain metal is to be used as a constructional material in industry, there must be a certain stock of various sizes available, from which to draw. It is little wonder that no use can be made — at least not for general factory use — of the hundreds of known aluminium alloys, not to mention the iron and copper alloys of which even more have been described, and the stream of new combinations in which practically all metals can be alloyed. A special alloy can certainly be taken for a special purpose, for instance, cadmium copper for the overhead conductors for electric trains; but for general use, many alloys are not directly available and a compromise is effected by making use as far as possible of those that are in stock. This situation can now be regarded differently. For making castings we can start from the *elements* — or from a limited number of master alloys. A casting can therefore easily be given the composition that best meets the requirements. This means that many alloys that have been developed during years of metallurgical work, but which could not easily be put to general use, can now be exploited. Many alloys with special mechanical, physical, magnetic or electric properties are now available for general engineering.

and a great deal of expense can often be saved by making a casting in such a way that it can take the place of a combination of machine-made sub-parts. This is especially so since extra-high demands are made for each of the separately machine-made sub-parts, in order to ensure that the total tolerances are not too great on assembly. (The total tolerances are equal to the sum of the tolerances of the component parts.) It is often cheaper to make one intricate casting and finish it by machining to the required final dimensions.

A second important advantage of the casting method described becomes evident when articles

An instructive example, in which not only the advantage just discussed, but also the "precision" of lost-wax casting plays a part, is seen in the employment of the stellites: hard, very wear- and corrosion- resistant alloys that can be machined only by grinding, which were developed and described about

1910 by Haynes and others [9]). Hitherto it had not been possible to make much use of these, because pieces cast in this material were so rough and inaccurate that machining became far too expensive. By application of the lost-wax casting process, a great future awaits them, as machine parts cast in this way from stellite need practically no machining.

To sum up: it may be stated that the lost wax casting method opens up the following possibilities:

1) The making of small but intricate castings in large or small numbers. In many cases, parts of production machines can be made that can combine various functions in one piece. For metals with a low melting-point, this method is rivalled by pressure die casting for large numbers (i.e., runs of more than 2000), provided the shape is not too complicated; for metals with a high melting-point the Croning process (shell-mould casting) is useful for turning out simple articles in large numbers.

2) It can be said that practically any alloy that is fusible can now also be used for small castings. As casting eliminates the necessity of keeping a wide range of metals and alloys in stock, attention can be paid to selecting those alloys that best meet the requirements of the work. These requirements may be of a mechanical nature (resistance to wear, corrosion and heat, and hardness) or of a physical nature (electric and magnetic properties, thermal expansion, thermal conductivity).

[9]) E. Haynes, Amer. Pat. 873745, December 1907 (!).
E. Haynes, Alloys of cobalt with chromium and other metals, Ind. Eng. Chem. 5, 189-191, 1913.
W. Oertel and E. Pakulla, Beitrag zur Frage der Kobalt-, Chrom-, Wolfram- (Molybdän)- Legierungen, Stahl und Eisen **44**, 1717-1720, 1924.

Summary. After a brief survey of the development of metal casting, and a description of sand casting and casting in permanent metal moulds (gravity die casting and pressure die casting), two modern processes for the casting of small objects are described, namely the Croning process (shell-mould casting) and the lost-wax process ("precision casting"). A number of details are considered of this last mentioned process, which has been investigated at Eindhoven and used for some years. The process is actually a very ancient one, but has found its way into industry only during the last decade. It is very suitable for making intricate castings and for the employment of many alloys with special properties, hitherto impracticable to use on a factory scale. Added to this is the advantage that the accuracy of dimensions and the quality of the surface of the cast parts are considerably better than, for example, in the case of normal sand casting.

BOOK REVIEWS

Television, by F. Kerkhof and W. Werner, pp. 475, 360 illustrations, 36 photographs — Philips Technical Library — This book can be ordered through your technical bookseller.

Writing a book on television is a rather hazardous enterprise, for in various aspects television technique is still undergoing rapid development. There is a real danger that some parts of a book on this subject will quickly become out of date.

The authors have been well aware of this danger and therefore aimed at focussing attention as much as possible on the physical principles, for these will remain the same — the chance that future television apparatus may be built on other principles is remote. The authors have succeeded in writing a book of lasting value, which will still be useful even when television techniques undergo further development.

The authors' intention has been to address themselves to as wide a group of readers as possible, i.e., to persons of varying education ranging from radio electricians to university graduates. In the opinion of the reviewer, the authors have succeeded in their aim.

The text can be easily followed. The essential properties of circuits are set forth. In many instances, calculations have been inserted in small characters for the benefit of those trained in mathematics; these calculations serve sometimes as the mathematical basis for the main text and sometimes to establish the circuit parameters more precisely. These passages can be skipped without making it difficult to understand the whole.

When reading the book, it becomes clear that the problems of television reception have been gone into more extensively than those on studio and transmitting apparatus. This is not surprising, since both authors are concerned with television receiver development at Philips. The consequence is that the book is most up to date in those chapters in which reception is discussed; this does not mean, however, that the studio and transmitting apparatus have been neglected. The discussion on these subjects is more simplified and the various associated problems are discussed less exhaustively.

A short survey of the subject-matter is as follows. The first chapter gives a general survey of the method in which the image is transmitted. In the second chapter the behaviour of electrons in magnetic and electric fields is described, while the third chapter treats camera-tubes and picture-tubes.

In the next chapter various transmitting systems are described and an analysis is made as to how the various "informations" needed in television transmission are transmitted and separated at the receiving end. The two chapters which follow, contain discussions on relaxation-oscillators and on deflection-generators. A whole chapter is devoted to H.T. generation for the picture-tube.

The problems relating to the amplification of wide frequency bands are discussed extensively. They are followed by treatments of cable-transmission and antennae, in two different chapters. The next chapter contains a description of the various optical problems which arise in projection television receivers. There is also a chapter on colour television, in which the various systems proposed in the USA by C.B.S. and R.C.A. are discussed. Finally, there is a chapter containing two complete diagrams of television receivers, one for a transmitting system with positive modulation, and one for a transmitting system with negative modulation.

At the end of the book are some appendices, including a list of definitions of terms used in television technique and an extensive bibliography. Sixteen photographs illustrate the most common faults in tuning a TV receiver; the appropriate remedial measures are also given.

The book forms a valuable addition to contemporary television literature. This is particularly the case for the chapters treating modern reception techniques.

J. HAANTJES.

Television Receiver Design, Monograph 1: I.F. Stages, by A. G. W. Uitjens, pp. 177, 114 illustrations. Philips Technical Library, Volume VIII A, Electronic Tubes series, 1953.

This book is the first volume of a specialist series, in which the various problems of television receivers design will be discussed. This series of books in the Philips Technical Library will form a welcome addition to the existing literature. The present volume deals with the intermediate-frequency amplifier of a television superhet receiver. The subject matter is, of course, equally applicable to the high-frequency amplifier of non-superhet receivers.

As may be expected in a book of the Electronic Tube series, the characteristics of the amplifier valves used are extensively dealt with.

In the first chapter the product *GB* (product of stage gain and band width) is introduced, and the correction factors to be applied in practical circuits are deduced.

Two-pole interstage coupling impedances are discussed in the second chapter. The reader is introduced to the cascade coupling of simple *L-C* circuits tuned to the same frequency and to amplifiers with *L-C* circuits tuned to slightly differing frequencies (staggered circuits).

In the third chapter the influence of frequency and phase characteristics on the step function response is discussed. The treatment includes not only symmetric side-band systems but also the asymmetric sideband systems generally used in television.

Four-pole interstage couplings are described in Chapt. IV; author introduces the term, "transfer admittance", which permits the use of the same mathematical treatment as for two-pole coupling impedances.

A very important feature of an amplifier is the signal-to-noise ratio. More than 20 pages of the fifth chapter are devoted to this subject. The various noise sources are discussed and some methods of calculating the noise factor are presented. The means by which this ratio may be improved is also discussed.

The author devotes the thirty odd pages of the sixth chapter to the discussion of the effects of stray couplings in the interstage networks and in the valves. The measures taken to minimize these effects are outlined.

The final chapter deals with the design of an I.F. amplifier with staggered circuits, making use of the formulae given in the previous chapters.

A number of appendices contain both the derivation of certain important general formulae and some convenient tables giving valve and circuit data.

W. WERNER

THE TROPOSPHERE AS A MEDIUM FOR THE PROPAGATION OF RADIO WAVES - I

by H. BREMMER.

621.396.11:551.510.52

*The hypothesis of Kennelly and Heaviside (1902) that an ionized layer existed in the atmosphere, which reflected radio waves and in this way explained satisfactorily why these waves are able to follow the earth's curvature — was confirmed experimentally in 1924 by Appleton. His experiments may thus be regarded as the beginning of research on the ionosphere. This has proved to be very important both theoretically and practically and it is now being pursued vigourously in many places all over the world. The methods used in this field of research have been reviewed in this journal by Prof. C. J. Bakker *).*

*Now that waves shorter than 10 metres are so widely used (for television, radar, etc.) it has become clear that for the propagation of these "microwaves" **), it is not the ionosphere that is important but the lowest layer of the atmosphere. This layer extends from the earth's surface to about 10 km above it, and is called the troposphere. Its behaviour in respect of radio propagation is closely related to meteorological conditions.*

Research on the troposphere is in full swing. Although it has not yet achieved a satisfying whole, as is the case for ionospheric research, it has gone far enough for us to be able to give a review of what is now known, especially in relation to radio wave propagation. The first part of this review follows.

Introduction

The simplest model for describing the propagation of radio waves around the earth is a perfect sphere surrounded by a uniform atmosphere. Radio waves can then only be propagated by means of so-called ground-waves. Even in this simple case, the calculation of the electromagnetic field is a complex problem. (It has already been discussed in this journal [1]). The results show that even for the shortest wavelengths used in practice — centimetre waves — the ground-waves extend to beyond the horizon of the transmitter [2]), and their range increases with the wavelength, but except for very long waves it is much less than the actual transmission range attained in practice. This is a consequence of the fact that all waves longer than about 10 metres are refracted in the ionosphere, between 70 and 400 kilometres above the earth's surface, and so may be deflected back towards the earth. However, the still shorter waves which have become extremely important for radar, television and transmitters with frequency modulation, are unaffected by the ionosphere. Instead they are refracted in the *troposphere:* any such waves reaching a receiver on the earth's surface have travelled exclusively through the troposphere. This is the name given by meteorologists to the lowest layers of the atmosphere, where the temperature decreases with height, down to a minimum of about —50° C at the "tropopause", at an altitude of about 10 km.

Microwave transmission can be effected only via the troposphere. In this respect microwaves behave like medium waves (about 200 m $<\lambda<$ 1000 m) during daylight. The ionosphere absorbs the medium waves during daylight, whilst it is always transparent to microwaves, so that in both cases, the troposphere is the only medium through which communication is possible.

In spite of this similarity, microwaves differ in behaviour from medium waves in two important respects.

1) Microwaves are much more closely related to optical waves, and like these, have only low intensities in shadow regions, such as beyond the horizon or behind a mountain. These are regions that the energy cannot reach if propagated according to geometrical optics, but can do so by diffraction. It would be exceptional, however, for the range of microwaves to extend to double the distance from the transmitter to its horizon.

2) Microwaves are far more sensitive to irregularities in the troposphere. This is understandable, for these irregularities are now more likely to be of the same order of magnitude as the wavelengths. In contrast, the behaviour of the longer waves depends rather on the average value, over one wavelength, of the irregularities, and is therefore less sensitive to them.

*) C. J. Bakker, Radio investigation of the ionosphere, Philips tech. Rev. 8, 111-120, 1946.

**) For the purposes of this article, the usual interpretation of the term "microwaves" has been somewhat extended to include all waves of $\lambda <$ 10 m.

[1]) B. van der Pol and H. Bremmer, The propagation of wireless waves round the earth, Philips tech. Rev. 4, 245-253, 1939.

[2]) Called "the horizon" throughout, for short.

Irregularities in the troposphere are always important in the propagation of microwaves. It is therefore not surprising that the behaviour of these waves is closely bound up with meteorological conditions. Their behaviour depends both on mean climatic or seasonal variations over a long time, and on short time variations (those of synoptic meteorology). Hence the properties of microwaves are not so easily surveyed as those of longer waves, for which the distribution of land and sea, for example, and daily fluctuations of temperature and humidity are not so significant in their effects. Many observations on both commercial and experimental transmissions have been made in order to gain more insight into these effects. It now seems clear that fluctuations in the atmosphere over minute distances — of the order of centimetres — are important, with which meteorologists were little concerned until very recently. Thus the investigation of microwaves has stimulated new meteorological research.

It is proposed to review here the theories put forward to explain the widely varying propagation characteristics of microwaves. The physical properties of the troposphere as medium for microwave propagation are first discussed.

The troposphere as a dielectric

The only physical quantity that is important for radio propagation is the refractive index n, which can equally well be written as $\sqrt{\varepsilon}$, where ε is the dielectric constant with respect to a vacuum, as conduction currents in the troposphere are negligible. We shall not yet consider how n varies from point to point, but only explain how n can be calculated from the physical properties of the gases in the troposphere.

The theory allows for the molecular polarization of each of the gases under the influence of electromagnetic waves. The theory states that not n itself, but $(n^2-1)/(n^2+2)$ is additive, i.e., the refractive index n_m of the mixture is given by:

$$\frac{n^2_m-1}{n^2_m+2}=C_1\frac{n_1{}^2-1}{n_1{}^2+2}+C_2\frac{n_2{}^2-1}{n_2{}^2+2}+\ldots$$

where C_1, C_2, ... are the volume concentrations of the component gases ($C_1+C_2+\ldots=1$), having refractive indices n_1, n_2, ... Now, for each of the gases, n differs only slightly from 1; the factor $(n+1)/(n^2+2)$ is therefore almost constant. Hence, for the mixture, a good approximation to n is given simply by summing values of $n-1$, which is very small compared with 1. Moreover, it appears that for wavelengths $\lambda>3$ cm, absorption effects can be neglected for all important constituents of the atmosphere (nitrogen, oxygen and water vapour). The refractive index is then real, and it is found that for microwaves with $\lambda>3$ cm it can be represented by:

$$n=1+79\times10^{-6}\frac{p_a}{T}+0.38\frac{p_w}{T^2}. \quad \ldots \quad (1)$$

Here T is the absolute temperature, and p_a and p_w are the partial pressures in millibars of dry air and water vapour respectively. This formula was derived theoretically and has been verified experimentally.

It is worth noting that the refractive index is rather sensitive to the amount of water vapour present, since the dielectric constant of water vapour is relatively high — the strong polarizability of the molecules of water falls off rapidly only at wavelengths shorter than those of microwaves. Further we note from (1) that n is independent of λ, so that the troposphere, unlike the ionosphere, causes no dispersion.

For wavelengths shorter than 3 cm, deviations from equation (1) begin to come in through absorption effects. Ultimately the absorption renders the use of very short wavelengths impossible for telecommunication, even though it is technically perfectly practicable to produce such waves.

For decreasing wavelengths, the absorption first perceptible is due to water (as a liquid) present in the atmosphere as rain, cloud or mist. The disturbance due to the water droplets can be calculated from the scattering and absorption in individual droplets. It is not so large as for an equal mass of water vapour. Both the absorption and the scattering can be derived, to a good approximation, from the dipole moment that is induced in each droplet by the field of the radio waves reaching it.

For water droplets in clouds and mist the calculations can be greatly simplified, for the droplets (max. diameter about 0.2 mm) are small compared with the wavelength. The theory is then the same as that for the scattering and absorption of light due to inhomogeneities in the atmosphere, whereby Rayleigh explained the blue of the sky. The amount of scattering depends mainly upon the *total volume* of the inhomogeneities, and not on their distribution as individual droplets, always provided they are all small compared with the wavelength.

Unlike light waves, radio waves are attenuated more by *absorption* in water droplets than by scattering. This absorption appears to be less than that for a coherent mass of water of equal volume.

When centimetre waves are absorbed by raindrops, the absorption is no longer independent of the size distribution of the drops, for they are not then small enough compared with the wavelengths. When this distribution is known, the absorption can still be calculated, although less simply. Rain can disturb the transmission of wavelengths shorter than 3 cm. For longer wavelengths than this, hail and snow absorb much less than the same amount of rain, but this is not necessarily so for shorter wavelengths.

For $\lambda < 2$ cm there is an appreciable absorption not only by water but also by the gases of the atmosphere. There is a resonance band near $\lambda = 0.5$ cm, and a resonance line near $\lambda = 0.25$ cm, both due to oxygen, and water vapour shows a single line close to $\lambda = 1$ cm. Both the band and the lines have been calculated theoretically [3]) and demonstrated experimentally. The propagation of the still shorter millimetre waves is rendered well-nigh impossible by strong absorption lines of water-vapour, but of course they can be used for communication over very short distances. The resonance band of oxygen has even been exploited to prevent a message sent over a short distance from being overheard by unauthorized persons further away.

In what follows, only waves longer than 3 cm will be discussed. For these waves the refractive index is given accurately enough by the simple equation (1).

Survey of idealized cases of radio propagation through the troposphere

The all-important refractive index at any point in the atmosphere depends, broadly speaking, on the altitude. The average values of n at different heights depend on the season and are closely related to the climate. At any one time and place, the change with height deviates from its mean value owing to the day-to-day changes in the weather, and these cause the index to vary from place to place at any one height. The result is that widely varying situations are possible, and these can influence the wave propagation to the good or the bad. We can simplify our description of what happens by enumerating several idealized types of wave propagation. Actually the waves will always be propagated in several ways at the same time.

In all these cases except the last, it is assumed that the refractive index depends only on the height.

[3]) See for example J. H. van Vleck, The absorption of microwaves by oxygen, Phys. Rev. 71, 413-424, 1947, and The absorption of microwaves by uncondensed water vapor, Phys. Rev. 71, 425-433, 1947.

The waves, then, may reach the receiver from the transmitter in the following ways:

a) By paths determined by laws of optical reflection at the earth and refraction in the atmosphere (*fig. 1a*).

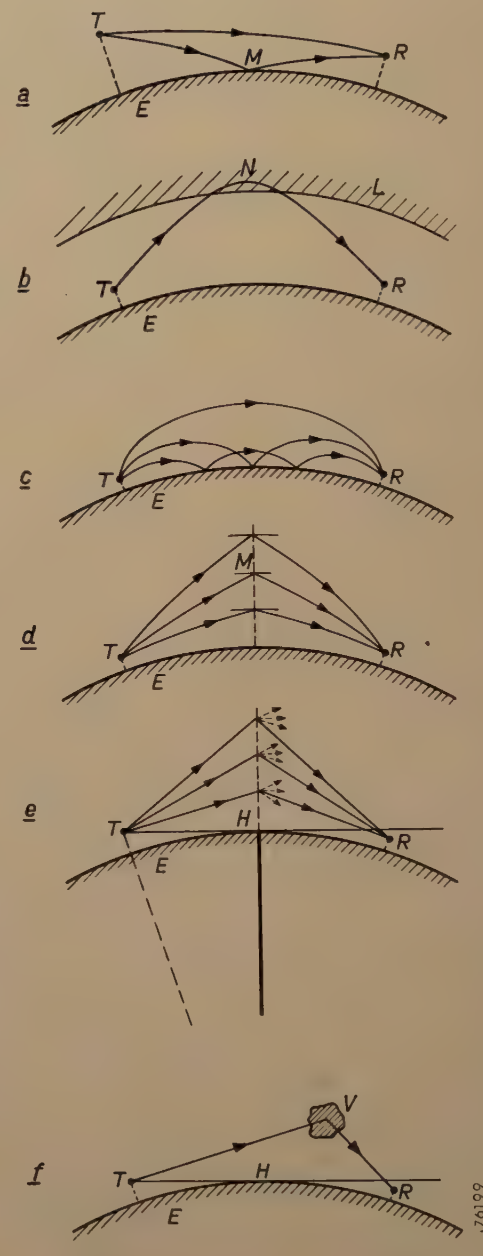


Fig. 1. Schematic representation of the different ways in which microwaves can reach a receiver R. T transmitter. E earth (curvature relatively greatly exaggerated).
a) Propagation along two paths, one direct (TR) and one reflected from the earth's surface.
b) Rays being continuously deflected in a transition layer (L).
c) Propagation by very many rays, as in a waveguide (super refraction).
d) Propagation via gradient reflections.
e) Although the receiver is below the horizon (H), rays reach R by diffraction, for example at the vertical plane through H.
f) R again below the horizon: rays reach R as a result of scattering at an inhomogeneity V in the troposphere.

b) By paths determined by the laws of geometrical optics, but occurring only under exceptional meteorological conditions (transition layers, fig. 1*b*); only a limited number of rays can reach the receiver.

c) By paths also determined by a transition layer, but which is due to other unusual conditions, very many rays may end at the receiver: they are propagated as if in a waveguide (fig. 1*c*).

d) By reflection at the thin "layers" of which the troposphere is composed, owing to its continuously changing refractive index (gradient reflections, fig. 1*d*).

e) By such paths that the field at the receiving end can be accounted for only by diffraction (fig. 1*e*).

f) By scattering at inhomogeneities in the troposphere, i.e., short-distance fluctuations in the refractive index both in a vertical and in a horizontal direction (fig. 1*f*).

Before discussing these cases in full, a few more general comments may be made. The transition layers in (b) and (c) may be regarded as surfaces of discontinuity if the refractive index changes within a height-interval less than one wavelength; reflection then occurs. In the more common case of a transition layer thicker than one wavelength, a ray progressing upwards (or downwards) can be bent continuously throughout its path within the layer, and so much so that it leaves the layer on the same side (*fig.* 2). In this case we speak of

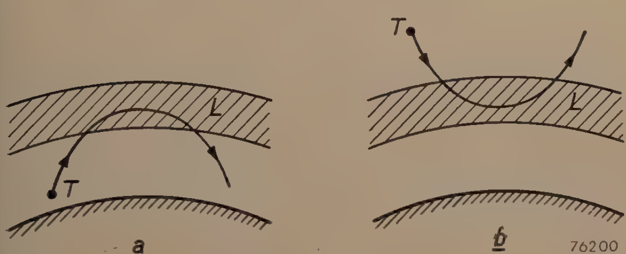


Fig. 2*a*) "Continuous reflection" in a transition layer *L*; rays reaching the layer from a transmitter *T* below *L* are bent back towards the earth.
b) *T* is above the layer *L*, which now bends the waves away from the earth.

continuous reflection at a layer, although the effect is in fact produced by refraction. Similar reflections occur when short waves are deflected from the ionosphere, and when sound travels a great distance (the sound waves are deflected at the ozone layer (ozonosphere), at an altitude of about 30 km).

An important question is — how large is the horizontal area over which the refraction index must be constant if it is not to disturb the paths based on a refractive index varying only with altitude? Suppose that the paths between transmitter *T* and receiver *R* are determined by geometrical optics. From the wave theory, whether or not a disturbance is caused by a local deviation in n (at P, say) depends on the path-length TPR; a disturbance is caused only if this path-length differs from the direct path (not necessarily straight) by less than $\frac{1}{2}\lambda$. The points P that satisfy this condition, lie within the "first Fresnel zone" with respect to TR.

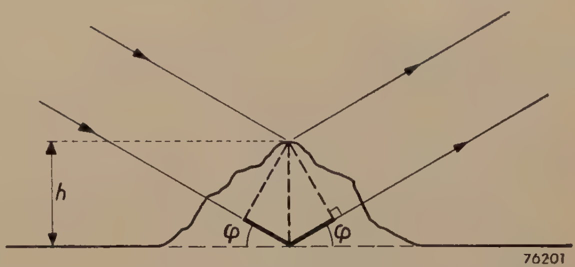


Fig. 3. The heavy line is the path difference between rays reflected at the top and the foot of the hill respectively, equal to $2h \sin\varphi$, when φ is the angle between the rays and the surface.

A geometrical study shows that if the maximum height of the ray is small compared with the distance D between T and R, the zone is concentrated within a narrow region about the vertical plane through TR; it extends only a distance $\frac{1}{2}\sqrt{D\lambda}$ from it. Thus, at points further from this plane, fluctuations in the refractive index are unimportant: for D = 100 km, for example, and λ = 5 cm, the critical distance is only 354 metres.

This narrowness of the first Fresnel zone is important, too, in relation to rays reflected at the earth's surface. In unbroken country we can regard the reflection as occurring on a plane, provided that over this plane the properties of the ground are uniform within the Fresnel zone. This zone, again, is a narrow one along the shortest path between T and R along the earth's surface, and extends for distances of about $\sqrt{D\lambda}$ on either side of the path. However, suppose a hill or other object is near the line TR, with a height h (*fig. 3*). Rays reflected from its foot and top respectively have a path difference equal to $2h \sin\varphi$, where φ is the angle of reflection (i.e. between the rays and the earth's surface). If this path difference exceeds the order of $\frac{1}{2}\lambda$, i.e. $\lambda < 4h\sin\varphi$, the reflection will be considerably modified — usually diminished.

a) Communication along the two normal paths

Whenever the receiver is above the horizon of the transmitter, it receives rays from both normal paths — the direct ray and the ray reflected at the

earth's surface. Both are curved since the troposphere is not homogeneous. If we now allow for the curvature of the earth and still assume that the refractive index n depends solely on the altitude, then n is a function of the distance r to the centre of the earth. It then follows from Snel's law of refraction that the product $r\,n(r)\,\sin\tau\,(r)$ is constant along each ray, where τ is the angle between the tangent to the ray at the point in question and the vertical through that point (*fig. 4*).

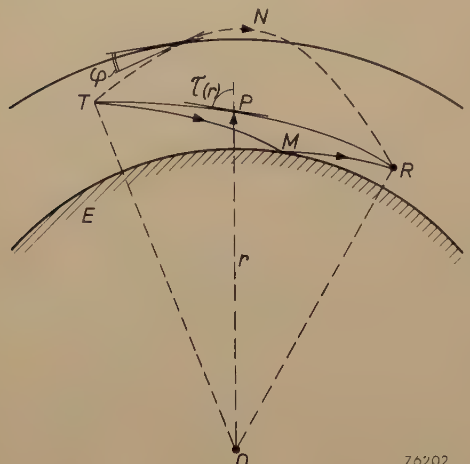


Fig. 4. Rays corresponding to fig. 1*a* (full curve) and fig. 1*b* (dashed curve). r distance between an arbitrary point P on the path and the centre of the earth. $\tau(r)$ angle between the tangent to the path at P and the vertical through P. φ = angle between an incident ray and the transition layer L.

Under normal conditions, $n(r)$ decreases with increasing height much more slowly than r increases, so that r dominates in the product $r\,n(r)$. According to Snel's law, therefore the angle $\tau(r)$ decreases along a rising path, and increases along a falling one. This is illustrated schematically in fig. 4. The propagation along *TR* is very much as it would be in free space, although the path is slightly curved. The field strength at R is therefore about inversely proportional to the length along the curve, and if the distance D along the earth's surface is much greater than the altitudes of the transmitter and receiver it is also inversely proportional to D.

The curvature of the shells with constant n, which are of course concentric to the earth's surface, causes slight deviations from this simple rule, and these make the paths diverge less for downward rays and more for upward ones, than would be the case in a homogeneous troposphere.

The field along the ray reflected at the earth, *TMR*, is also about inversely proportional to the distance along the path from T, if we ignore the effect on the intensity of the reflection at M. For all but extremely short waves, this reflection, too, can be regarded as coming from a flat surface (see end of previous section). The corresponding reflection coefficient can be derived from Fresnel's formula, which is based on Maxwell's equations. For microwaves, this reflection coefficient proves to be very close to -1, since the waves almost graze the earth's surface. This means that the reflection is as good as total, but it is accompanied by a phase change of close to π radians. However, for waves shorter than 10 cm, the earth's surface is not usually smooth enough to be regarded as flat, and the theory can no longer be applied. The reflection is more like scattering at a rough surface, giving an effective reflection coefficient of the order of 0.2.

The total field intensity at the receiver is found by adding together the fields produced by *TR* and *TMR* and allowing for their difference in phase. The latter is partly due to the difference between the two path lengths, and partly to the phase shift of nearly π where the second ray is reflected (when this can be regarded as reflected at a flat surface). The phase difference means that, at a fixed height, the field intensities vary with distance. These interference effects are of course much weaker when the reflecting surface at M is "rough", for then the field due to *TMR* is quantitatively unimportant.

b) Communication via transition layers

The presence of a large gradient of the refractive index in the transition layers is contingent on the meteorological conditions. Such a layer is usually accompanied by an inversion (i.e. the temperature *rises* with height), or at least by a much smaller fall of temperature with height (lapse rate) than the average. It is not so much the temperature itself, but the corresponding changes in pressure of water vapour that are mainly responsible for an abnormal change in the refractive index (see equation 1). These transition layers are found especially at heights between 500 and 2000 metres, and the change in n across the layers is often of the order of 5×10^{-6} (while $n-1$ is round about 3×10^{-4}, see above). Slowly-moving transition layers often occur in settled weather; on the other hand, layers may occur that extend along meteorological fronts and move with them.

When transition layers are present, propagation along paths like the dashed line *TNR* in fig. 4 is possible. This kind of path is especially important when the normal paths *TR* and *TMR* are impracticable because the receiver R is below the horizon; apart from this, R can be reached only by diffraction, and the corresponding field at R is then weak. In

particular, it has been established that the field is temporarily strengthened when a meteorological front passes between transmitter and receiver. Several transition layers can be operating at the same time [4]).

The field along a path via a transition layer is, once again, about inversely proportional to the distance from the transmitter, but the effect of reflection at the layer must be allowed for. Two limiting cases will be discussed in which it is easy to give a general idea of this effect: these are, when the layer is much *thinner* than the wavelength and much *thicker*.

Layer thickness small compared with the wavelength

For thin layers the conditions are nearly those of reflection at a surface of discontinuity of the refractive index, so that the coefficient of reflection may be calculated from Fresnel's theory. The angle of incidence φ on the layer (see fig. 4) is small, yet it is larger than the critical angle (below which the reflection is total), and because of this the reflection is partial: part of the energy of the incident ray gets through to the space above the layer. Further, the angle φ is small compared with Brewster's angle (that is the angle for which there is zero reflection of waves with electric vectors in the vertical plane through the incident ray). Because of this, a good and simple approximation to the reflection coefficient R proves to be

$$R \approx \frac{\Delta n}{2n \sin^2\varphi}.$$

Here Δn is the difference between the refractive indices on the two sides of the layer. Thus the reflection coefficient does not depend on how n changes inside the thin layer. When the distance D is large compared with the height h of the layer above the earth, φ and $\sin\varphi$ can both be replaced by $2h/D$, and then

$$R \approx \frac{\Delta n}{8n} \cdot \frac{D^2}{h^2} \quad \ldots\ldots \quad (2)$$

from which it is clear that a given layer reflects best when the transmission is over great distances. This is understandable, for then the incident angles may approach the critical angle for total reflection.

Layer thickness large compared with the wavelength

The reflection at thick layers can be said to be "continuous" and appears to be total, i.e. no rays pass through the layer. However, according to the theory, the continuous reflections are associated with an extra phase shift of $-\frac{1}{2}\pi$, which can be compared with similar phase shifts in optics.

Actually, all intermediate layers between the two extreme ones are possible. Their reflection coefficients have been calculated by making assumptions about the way n changes through the layer, and in this way Epstein [5]) and Rawer [6]) have demonstrated that the reflection is generally better for the longer wavelengths.

The fields due to reflections at one or more transition layers must be compounded vectorially with the field due to the two normal rays (discussed under *a*), provided of course these exist, i.e. the receiver is above the horizon. What is observed is often so complex that it is difficult to say whether or not transition layers play any part. At great distances they can be recognized by the increase of the reflection coefficient with distance under eq. (2), which can partly compensate for the normal weakening of the field. Apart from this, the transition layer may be revealed by an increase of the field for the longer wavelengths, due to the greater reflection coefficient of the layer, but this can be masked by less reflection of the same wavelengths at the earth's surface. Total reflections at transition layers can be observed only for low-altitude layers or by observations from the air, in which cases the angles of incidence φ can be very small.

c) Cases where the troposphere acts as a waveguide

In general, ascending waves escape into space, but this does not necessarily happen to all rays. For example, some rays may be deflected back by continuous reflection, and then remain confined to the troposphere by reflection between the earth's surface and the highest "continuous reflecting" surface. The space in which the rays travel to and fro between the earth's surface and the continuous reflecting surface has all the typical properties of a waveguide [7]). The theory of this effect was given by Booker. The conditions for the existence of such a waveguide can be derived as follows. A ray undergoing continuous reflection must attain a maximum height somewhere, and there its tangent is horizontal and τ is 90°. Since by Snel's law, $rn \sin\tau$ is constant, τ can become 90°

[4]) For a survey of the effect of transition layers, see an article by J. A. Saxton and others, Proc. Inst. El. Engrs. **98**, III, 360-378, 1951.

[5]) Proc. Nat. Acad. Sci. **16**, 627-637, 1930.

[6]) Ann. Physik **35**, 402, 1939.

[7]) For the properties of waveguides, see e.g. articles by W. Opechowski in Philips tech. Rev. **10**, 13-25 and 46-54, 1948.

only at that point on the path where rn has the smallest value. There must therefore be a region where rn decreases over a rising path. Such a region is called a duct, and it is not often met with, for usually r increases more than n decreases. In the exceptional cases where n decreases strongly enough for rn to decrease and produce a duct, we speak of super-refraction.

The presence of a duct does not mean that *all* rays going upwards are bent downwards again by continuous reflection. To consider a ray that is deflected back, let r_0 and n_0 be respectively the values of the distance to the earth's centre and the refractive index at the highest point in its path, and a and n_a the corresponding values at the earth's surface. Then for a ray leaving the earth's surface at an inclination τ_a to the vertical, Snel's law gives

$$a\, n_a \sin \tau_a = r_0 n_0.$$

When rn is a minimum (denoted by $(rn)_{\min}$), i.e. on the upper side of the duct, τ_l, the inclination, to the vertical of the steepest ray to be reflected, is given by

$$a\, n_a \sin \tau_l = (rn)_{\min}.$$

A similar limiting angle is also known in propagation through the ionosphere. Only the less steeply rising rays can be continuously reflected back towards the earth from within the ionosphere. But continuous refraction in the ionosphere and super-refraction in the troposphere differ in several respects, as follows:

1) In the ionosphere, continuous reflection and a critical angle for it are quite normal, whilst in the troposphere, super-refraction only occurs under exceptional meteorological conditions [8]). Over the sea, conditions are often favourable for super-refraction, due to the relative humidity falling off quickly with height. Over the land, the cooling of the surface at night is very favourable for super-refraction, when a layer of cold air is formed over the ground. Here also the decrease of humidity with increasing height is very important: in dry air super-refraction is only possible in a temperature inversion with rise of more than 11.2 °C in 100 m, which is highly improbable, whilst a much smaller temperature gradient suffices in damp air. (In contrast to this, acoustic super-refraction can occur with nearly every temperature inversion.)

2) In the ionosphere, because of dispersion, the limiting angle depends very much on the frequency, but this is not so for super-refraction.

3) In propagation via the ionosphere there is a dead zone (skip distance) around the transmitter, and rays can reach the earth through continuous reflection only outside this zone. The receiver R can be reached only by a few rays in this way. For super-refraction in the troposphere, on the other hand, when the duct lies just above the earth's surface, very many rays (in theory infinitely many) go from the transmitter to the receiver according to Snel's law. The rays differ in the number of reflections undergone. This great number of rays is possible because there is no homogeneous region between the earth and the duct — as there is effectively between the earth and the ionosphere. As a result, the distance from the transmitter to the point where the ray first returns to the earth's surface can decrease only when the angle τ_a increases (*fig. 5a*) — in contrast to the case of ionospheric reflection.

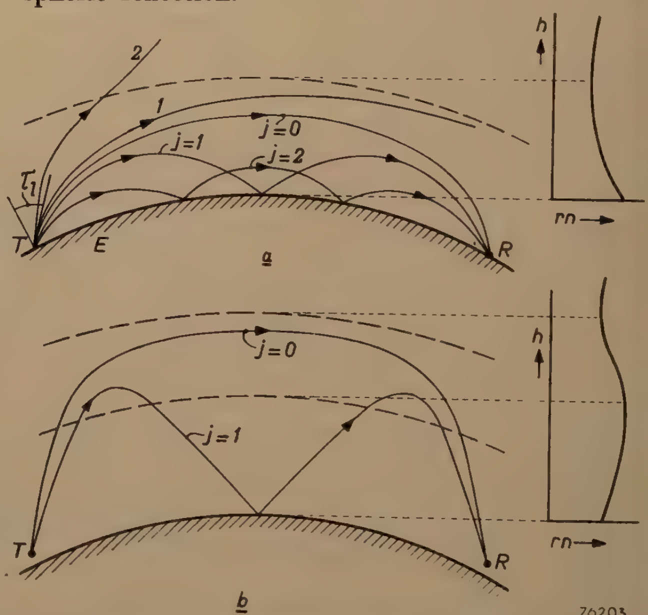


Fig. 5. Cases of super-refraction. There is sometimes a region in the troposphere ("duct") within which rn decreases with rising height, either (a) from the ground upwards ("surface duct"), or (b) only above a certain height ("elevated duct"). Many rays can now go from the transmitter to the receiver, and they undergo j reflections at the earth's surface and reach maximum heights $j+1$ times.

The ray *1*, sent out in the direction of the critical angle τ_l at T, has for its asymptote a circle concentric with the earth. The steeper ray *2* does not return to the earth.

On the right rn (horizontal) is drawn as a function of h (vertical).

The typical characteristics of a waveguide in the troposphere are not so pronounced when the duct, (that is the region where rn decreases with rising height) only begins well above the earth's surface. In such an elevated duct, the number of possible connecting rays is limited since the limiting angle

[8]) These conditions have been treated in detail, in an article by H. G. Booker, Elements of radiometeorology, J. Inst. El. Engrs. **13**, IIIa, 69-78, 1946.

τ_l is larger (fig. 5*b*); there is then only a vague distinction between an elevated duct and a transition layer producing total reflections. But for non-total reflection, the longer the wavelength the better the reflection at a transition layer, whereas for an elevated duct the contrary is the case.

So far only the geometrical properties of super-refraction have been discussed. Of these, the most notable is that the range is much larger than normal. Something analogous might be expected with light waves, but their absorption is generally too strong for super-refraction phenomena to be observed. Optical visibility can sometimes be greater than is normally possible, but it can never extend to several times the distance to the normal horizon, as radio waves can. We may mention that mirages (for example, the well-known "fata morgana") constitute a counterpart to super-refraction, as they are generally seen when the refractive index decreases with height less than normally (or even increases); such cases of under-refraction can also occur with radio waves (see section e, on diffraction).

In the space between the earth and the duct the ray paths described exist for all frequencies; however, as in a conventional waveguide, the interference of the fields due to the various rays can only lead to an abnormally strong field if their frequencies exceed a certain limiting one. In the next section we shall discuss the limiting frequency corresponding to a surface duct more fully. This is the simplest case, since the duct is not then merely one boundary of the waveguide, but coincides with it.

Existence of a limiting frequency for a tropospheric waveguide

The waveguide with which the troposphere may be compared when it is super-refracting has the peculiarity of having no firm wall for its upper boundary, whilst its lower boundary is always the earth's surface. The rôle of the upper boundary is played by the concentric sphere (to the earth) from which a certain ray can be continuously reflected, although the level of this sphere depends on which abnormal ray is being considered (fig. 5*a*). The vagueness of the upper boundary has the result that the theory of the tropospheric waveguide is much less simple than that of a normal waveguide bounded by solid walls. Despite this, the theory can be developed in two different well-known ways, namely by geometrical optics or from the complete wave equations.

Geometrical-optical explanation

In the geometrical-optical way of regarding the problem, the field at the receiver R is interpreted as the vector sum of infinitely many rays from T. The path lengths of the rays are all about the same, so their separate contributions would be about equal if we ignored the reflections at the earth's surface (j of them for each ray) and the continuous reflections above ($j+1$ in number). For both rays, both sorts of reflection are nearly total and the accompanying phase shifts are π and $-\frac{1}{2}\pi$ radians respectively (see above). From this it follows that although the amplitudes of the separate components are scarcely changed by the reflections, the mutual phase differences of the fields due to successive rays (for which the parameters j differ by unity, see fig. 5), become about $\frac{1}{2}\pi$; and it appears that for not-too-short wavelengths the differences between path lengths are less important. For each ray one can therefore always find a second ray with a j value differing by only a few units (and hence nearly enough the same amplitude), such that the second ray will more or less extinguish the first one. This interference effect weakens the resulting field very markedly. There is a way out of this difficulty, but only when the phase shift due to differences in path lengths can contribute an extra $\frac{3}{2}\pi$ for a number of successive paths. The total phase difference of $\frac{3}{2}\pi+\frac{1}{2}\pi$ for two consecutive rays then effectively corresponds to a whole wavelength. In practice this can only happen for very short waves, so that only from them can we expect to get a reasonable field strength. This explains in principle why a limiting frequency exists.

On the basis of this model the limiting frequency can be calculated if we assume a suitable function for the change of refractive index within the duct with height; the function must of course be consistent with the condition that rn decreases with height. From a very simple model (r^2n^2 a polynomial of second degree in $1/r$) the wavelength λ_0 corresponding to the limiting frequency comes to

$$\lambda_0 = \tfrac{4}{3}\,h_0\cos\tau_l = \frac{4\sqrt{2}}{3}\,h_0\sqrt{\frac{\delta(rn)}{rn}}. \quad . \quad (3)$$

Here h_0 is the altitude of the top of the duct, i.e. the height where rn is a minimum, and so, by the definition of the limiting angle τ_l, the height approached by a ray leaving the transmitter at the limiting angle; $\delta(rn)$ is the difference between the values of rn at the top of the duct and at the earth's surface.

Equation (3) explains numerically why the waveguide effect in the case of super-refraction is

perceptible only for very short wavelengths. The value of cos τ_l is very small in practice and h_0 is of the order of several tens of metres or several hundreds; hence λ_0 is usually in the region of centimetre waves. In any event, super-refraction was not discovered until the second world war, when centimetre waves were used on a large scale for the first time (to observe aircraft and ships by radar).

Equation (3) also demonstrates the consequences of having an indefinite upper boundary for the duct; the factor cos τ_l is a measure of the indefiniteness, for the factor would be unity if every ray directed upwards was bent back to the earth in one way or another. This does happen, more or less, with reflections of longer waves in the ionosphere, and in this case the whole space between the earth and the upper boundary of the ionosphere can be regarded as a wave guide. The limiting wavelength of this wave-guide is about equal to its width, that is roughly the height h_0 of the ionosphere above the earth. This height is of the order of 100 km, which is greater than any radio wavelengths ever used in practice, and that makes it clear why the ionosphere is a propagating medium for all radio waves long enough to be refracted by it.

Explanation by wave theory

The treatment of radio-propagation in the troposphere with a refractive index depending only on the altitude, generally amounts to the following, whether super-refraction is present or not. From the three-dimensional wave equation, solutions are determined that consist of products of a function of r only and a function of ϑ only, where r and ϑ are polar co-ordinates in a vertical plane through the transmitter with the centre of the earth as origin. $\vartheta = 0$ denotes the vertical through the transmitter (*fig. 6*). From the infinite

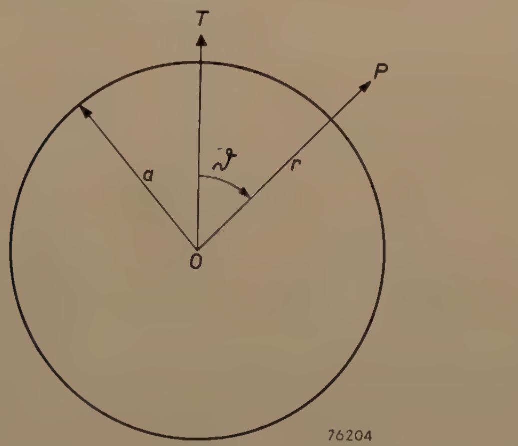


Fig. 6. Polar co-ordinates r and ϑ of a point P, with the centre of the earth as origin O. Angle ϑ measured from the vertical through the transmitter T.

number of solutions of this form for a given frequency, particular solutions are sought that are zero at infinity and satisfy the boundary conditions imposed at the earth's surface. These solutions constitute the modes, which are infinite in number but form a discrete assembly. Each mode represents a condition in which no energy is brought into the system from outside, while the field strength along the vertical through the transmitter is infinite. By choosing the right combination of these modes not only is a field obtained that is infinite at the transmitter — regarded as a point source — but also the singularity there can be chosen to agree with the known data concerning the transmitter (radiated power, polar diagram.)

The limiting frequency now comes into the theory in the following way. The ϑ-dependent factor in the mode-solutions of the wave equation is approximately proportional to

$$\frac{e^{a_r\vartheta}}{\sqrt{\sin\vartheta}},$$

where a_r is different for every mode. The exponents a_r are generally complex with a negative real part, so that for increasing angular distances the wave function falls off exponentially. Only in the case of super-refraction, and even then, only for high enough frequencies, are there one or more modes with purely imaginary values of a_r, corresponding to a wave-function not decreasing exponentially with ϑ. The limiting frequency is then the lowest frequency that has at least one of the latter modes; for them the amplitude of the field is proportional to $(\sin^{-1/2}\vartheta)$ and so decreases as ϑ increases even more slowly than for waves spreading out in a vacuum (where the field is inversely proportional to the distance, i.e. for short distances proportional to ϑ^{-1}). The duct action of super-refraction for frequencies above the limiting one can thus be explained in terms of the contribution to the field strength from modes with purely imaginary values of a_r: we shall denote them by $\mathrm{j}\beta_r$.

The wave function for these modes — each is then a function $(\sin^{-1/2}\vartheta)\exp(\mathrm{j}\beta_r\vartheta)$ multiplied by a function of r — may be studied more fully, and it is found that they are geometrically-optically equivalent to a system of curved rays, which, like the rays of fig. 5, travel up and down more than once and are alternately reflected discontinuously at the earth's surface and continuously in the troposphere. All rays belonging to one such mode remain in one beam and neither converge nor diverge, and are derived from one another by rotating the whole curve through an arbitrary

angle in the direction of ϑ. In fact these rays are mutually congruent (*fig.* 7). Further it can be deduced that the inclination $\tau_{a,r}$ to the vertical of the ray at the earth's surface is given by

$$\sin \tau_{a,r} = \frac{\lambda}{2\pi a}\,\beta_r \,. \quad . \quad . \quad . \quad . \quad (4)$$

What is especially significant about a system of rays and modes with $a_r = \mathrm{j}\beta_r$ is that energy is transported along these rays almost without loss; whilst it is much less simple to describe what happens

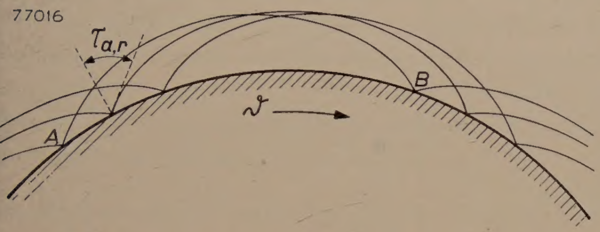


Fig. 7. In the case of super-refraction, the rays belonging to one mode, with $a_r = \mathrm{j}\beta_r(\beta_r$ real), can be obtained from one another by rotating the whole of a curve representing one ray about an arbitrary angle in the direction of ϑ. $\tau_{a,r}$ is the inclination of the ray to the vertical where it is reflected at the earth's surface.

when a_r is complex. In the latter case, the energy travelling outwards in the ϑ direction diverges, and part of it leaks away upwards and does not remain concentrated within a layer just above the earth's surface, as it does when a_r is purely imaginary. These two types of modes have been named "leaking modes" and "trapped modes" respectively.

Given the exponent $a_r = \mathrm{j}\beta_r$ for a trapped mode, equation (4) shows that a system of rays that can actually return to the earth is only possible if $\lambda < 2\pi a/\beta_r$. This establishes that each of these modes has its own limiting frequency, while the general limiting wavelength λ_0 then corresponds to the smallest value of β_r. The boundary conditions at the earth's surface impose severe restrictions on the possible values of β_r and hence of $\tau_{a,r}$, and this can be interpreted as a quantizing of directions for a given frequency. This must not be confused with the quantizing in many other problems involving characteristic functions, where unknown *frequencies* are quantized and not unknown directions. The direction quantization can also be regarded as arising from the condition that the total phase change between two similarly situated points A and B, due to a rising and falling ray path, must be a multiple of 2π.

To sum up the two treatments: using geometrical optics, we start with the transmitter and the rays sent out from it, and we are only later concerned with the boundary conditions at the earth's surface, when calculating the reflection coefficient. On the other hand, in the wave-theory method we start with the boundary conditions, and from these deduce the modes, and afterwards, when combining the modes, allow for the transmitter.

d) Gradient reflections

The transition layers discussed in a previous section were rather vaguely defined, in that we spoke only of a rapid change in the refractive index over a small interval of height. In fact, every infinitesimally thin spherical shell between heights h and $h+\mathrm{d}h$ can be regarded as a layer in which n deviates from its value in the adjoining layers, so that the layer can act as a partial reflector for rays coming from both below and above [9]). For a layer to reflect, it is only necessary that n should not vary much horizontally within the Fresnel zone corresponding to the reflection. Apart from this, these reflections exist in theory whenever the refractive index varies with the height; thus the gradient of n is the origin of the reflections, which may be termed gradient reflections. The ray changes its direction discontinuously, in contrast to what happens with the continuous deflections shown in fig. 1*b*.

In the field at the receiver, contributions are to be expected not only from single gradient reflections (as drawn in fig. 1*d*), but also from multiple gradient reflections, whereby the ray travels to and fro between the different shells several times. The quantitative significance of gradient reflections is still an object of discussion.

A feature of gradient reflections is that they too can contribute to the field far beyond the horizon. The receiver R can always be reached (in the case of *single* gradient reflections) by means of reflections in the plane of symmetry of the arc TR (fig. 1*d*), provided the distance TR is not so large that the point of reflection lies in too rarified air. In any case it is understandable that the differences in field strength obtained for homogeneous and inhomogeneous tropospheres (in which n depends only on the height), can be formally ascribed to gradient reflections, as Feinstein [10]) has shown. Usually, however, this variation in the field is explained by diffraction theory (see next section).

In general, the field resulting from gradient reflections is strongly diminished through the extinguishing interference between the contributions

---

[9]) H. Bremmer, Physica **15**, 593-608, 1949.
[10]) J. Feinstein, Trans. Inst. Rad. Engrs., Professional group on Antennas and Propagation, Dec. 1952, pp. 2-13.

arising from each of the extremely thin individual layers. These contributions reach the receiver with widely differing phases. Because of this it is assumed that fields observed far beyond the horizon that are too strong to be explained by diffraction, do not come from gradient reflections but from scattering (see Part II of this article). It has to be remembered that the variations in the refractive index, which in any case are strongest in the vertical direction but also exist horizontally, make it difficult to define a sharp distinction between gradient reflections and scattering. In the region above the horizon of the transmitter, the effect of gradient reflections is entirely drowned by the field due to the normal rays discussed above.

e) Fields due to diffraction

Under normal conditions (super-refraction absent) a point on the horizon of the transmitter is the point of contact of a ray just touching the earth. This ray is always bent to some extent through the ever-present refraction, usually downwards, as the refractive index normally increases with approach to the earth. The horizon therefore lies rather further away than it would do under a perfectly homogeneous troposphere. The region beyond the horizon beneath the tangential rays is inaccessible to geometrical-optical rays, if we ignore gradient reflections and transition layer effects (if any). The field in this region can be explained only by diffraction, in the same way as the penetration of light into the edges of shadows.

The deflecting field is usually investigated with the help of the method of the modes, already discussed. The wave functions of the modes were proportional to $(\sin^{-1/2}\vartheta)\exp(a_r\vartheta)$, wherein each exponent a_r has a negative real part $-c_r$ (leaking modes) if we ignore super-refraction. The decrease of the wave function at increasing angles from the transmitter is apparently mainly determined by the factor $\exp(-c_r\vartheta)$.

The field is given by the infinite series that combines all modes, and the series converges everywhere, including above the horizon, although there very slowly. The situation is simplest for distances so great that only the first term (that with the smallest c_r) of the series is numerically important. It is therefore this smallest c_r that determines the range of the transmitter. The most important c_r-values have been found numerically for many physical models, in which the dependence of n on the height is chosen according to meteorological data. A general theory for this has been given by Pekeris [11]).

The results of calculations on any one model for a non-homogeneous troposphere can now be compared with those for a homogeneous troposphere. For the latter, and a given state and type of ground, the field substantially depends only on the following dimensionless parameters:

$$\frac{D}{a^{\frac{2}{3}}\lambda^{\frac{1}{3}}},\ \frac{h_1}{a^{\frac{1}{3}}\lambda^{\frac{2}{3}}},\ \frac{h_2}{a^{\frac{1}{3}}\lambda^{\frac{2}{3}}},\ \ldots\ldots \tag{5}$$

where a is the radius of the earth, and h_1 and h_2 are the heights of the transmitter and the receiver respectively, above the ground. When the gradient of the refractive index near the earth's surface varies only slightly, it appears that the results for an inhomogeneous troposphere can be obtained from those for the homogeneous one by replacing a in the three parameters in (5) by an "effective radius":

$$a_{\text{eff}} = \left[\frac{r\times n(r)}{\dfrac{\mathrm{d}}{\mathrm{d}r}\{r\times n(r)\}}\right]_{r=a} = \frac{a}{1+a\dfrac{n'(a)}{n(a)}}. \tag{6}$$

Usually, the result of changing the parameters (5) is expressed in another way, namely that in the formulae for the field in a homogeneous troposphere the horizontal distances have to be multiplied by $(a/a_{\text{eff}})^{\frac{2}{3}}$ and the heights by $(a/a_{\text{eff}})^{\frac{1}{3}}$. Normally $n'(a)$ is negative, so that $a_{\text{eff}} > a$; the apparent distances are therefore smaller than the real ones. This agrees with the fact that under average atmospheric conditions, and when the transmitter and receiver are on the ground, a field is observed which is stronger than would occur if the troposphere were homogeneous.

One can obtain a picture of the effective earth-radius in geometrical-optical terms. Consider a ray that starts from a height $h = r - a$ and moves horizontally. The initial curvature of the ray is found to be

$$\frac{1}{\varrho} = -\frac{n'(r)}{n(r)}.$$

If $\varrho = r$, the ray has the same curvature r as the concentric sphere through its starting point, i.e. the ray "follows" the curvature of the earth. In this case, the ray remains at a constant height above the earth, as would also be the case with a *homogeneous* troposphere above a *flat* earth. It is understandable that in all other cases the field depends primarily upon the difference between the curvatures $1/\varrho$ and $1/r$. On the earth's surface, this difference is exactly equal to $1/a_{\text{eff}}$, whereby the

[11]) C. L. Pekeris, J. appl. Physics 17, 1108-1124, 1946.

significance of a_{eff} becomes evident. To obtain the field in hilly country, it is reasonable to replace the actual curvature of the earth by that of a circle that approximates as nearly as possible to the profile of the ground between the transmitter and the receiver.

It has been mentioned that the fall in temperature and humidity with height is, under most meteorological conditions, such that $a_{\mathrm{eff}} > a$; a_{eff} is often about $\frac{4}{3}a$. It can happen, however, that $a_{\mathrm{eff}} < a$, especially with a low hanging mist. The conditions for propagation are then worse than in a homogeneous troposphere.

Finally it may be remarked that, even for microwaves, the field never changes abruptly at the horizon for either a homogeneous or a non-homogeneous troposphere.

Limiting cases of very favourable conditions for propagation

To illustrate the conditions that govern the all-important variation of the refractive index, a few idealized limiting cases may be mentioned.

Using the conception "effective radius" of the earth, super-refraction is represented by a negative value of a_{eff}, for the differential coefficient $\partial(rn)/\partial r$ is negative in a duct. When a_{eff} changes its sign, it passes through infinity, and the product rn then has a stationary value at the earth's surface. When close to the surface, rn then varies only slightly with height, and properties not very different from those in an atmosphere with a constant rn can be expected.

In this kind of atmosphere it is easy to see that any one ray must have the same inclination τ to the vertical everywhere, in order to satisfy the path equation $rn \sin \tau =$ constant (Snel's law). In particular, a horizontal ray remains horizontal, and in theory goes round the earth and comes back to its starting point. In such a medium, one could see the back of one's head if light absorption did not exist! Fleming calculated quite some time ago [12]) that this condition would be very nearly realized if the earth was surrounded by an isothermal atmosphere of krypton with the same temperature and pressure at the earth's surface as the real atmosphere. (In such a krypton atmosphere the pressure would change with height differently of course from our atmosphere.) This is a nice illustration of how the molecular composition of the atmosphere is involved in the refraction index and its changes.

The situation with a constant value of the product rn is actually nearly achieved if super-refraction occurs, near the upper boundaries of the ducts ($r = r_0$); since such a boundary is characterized by a minimum value of rn. Near the level $r = r_0$, rn thus changes very little. Hence at this level, a horizontal ray does indeed remain "floating" and follows the curvature of the earth (ray *1* in fig. *5a*). This is certainly not the most favourable condition for propagation over great distances, for a ray very slightly inclined to the spherical surface $r = r_0$ and leaving the sphere will get a curvature that deflects it still further away. However, when rn is not a minimum but a maximum, rays slightly diverging will be deflected back to the direction parallel to the earth's surface; and this happens on the lower side of an elevated duct. A similar situation exists for acoustic refraction in seawater because the refractive index is influenced both by decreasing temperature and by increasing pressure at greater depths. Exceptionally favourable conditions for acoustic propagation in the sea are realized at a depth of about 1 km. At that depth the product rn (the factor r of which is scarcely important) goes through a maximum. The principle has been put to use in the so-called "Sofar" technique to detect very distant explosions of depth charges: the charges are set to go off at the depth of the duct, and microphones are placed at the same depth. In this way acoustic signals have been sent over many thousands of miles.

We have now discussed five of the six characteristic mechanisms of radio propagation. The sixth, scattering, will be the subject of the second part of this article.

[12]) J. A. Fleming, Proc. Phys. Soc. London, **26**, 328, 1914.

Summary. This article forms the first part of a review of what is known about the troposphere in relation to radio transmission. The most important physical magnitude that influences propagation of electromagnetic waves through the atmosphere is the refractive index n. A formula for n is given as a function of temperature and of the partial pressures of dry air and water vapour in the atmosphere, valid for waves longer than 3 cm. For shorter waves there are deviations due to absorption, by water droplets, and for waves shorter than 2 cm by water vapour and by oxygen. The rest of the discussions are confined to waves longer than 3 cm (and shorter than about 10 m).

Six idealized possibilities are distinguished (which can arise in many different combinations) whereby radio communication can be made using microwaves:

a) along two rays, one direct and one reflected at the earth's surface, both according to the laws of optics (the normal case),
b) along one or a few rays via a transition layer,
c) along many rays that are propagated as if in a wave guide (super-refraction),
d) via reflections brought about by a continuous change in the refractive index with altitude (gradient reflections),
e) by rays which are diffracted, and
f) by rays which suffer scattering.

Cases (*a*) to (*e*) have been considered in some detail. The appearance of a duct is discussed under (c), and the existence of a limiting frequency is explained. Scattering (case *f*) will be dealt with in the second part of the article.

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the address on the back cover.

2032: J. D. Fast: De invloed van verontreinigingen op de eigenschappen van metalen (Metalen 7, 2-12, 23-27, 48-50, 1952, Nos. 1, 2 and 3). (The influence of impurities on the properties of metals; in Dutch).

The deleterious influence of impurities on the properties of metals and alloys is discussed. Inclusions may, however, also have a favourable influence on certain mechanical properties of metals and alloys, by inhibiting the growth of crystals. Very useful effects (mechanical hardening, magnetic hardening and resistance against creep at high temperatures) may be obtained with inclusions formed by precipitation from a supersaturated solution. The same effects are deleterious if mechanically or magnetically soft materials are wanted. From a study of the effects of carbon and nitrogen in pure iron, and in iron containing 0.5 per cent manganese, it is possible to give an explanation of the phenomenon of magnetic ageing of steel. The interaction between impurities and dislocations is discussed in connection with Cottrell's theory. Several facts in the domain of strain ageing can thus be explained, but other facts do not fit very well in this theory. (See also Philips tech. Rev. **14**, 60-67, 1952).

2033: F. A. Kröger, H. J. Vink and J. van den Boomgaard: Absorption and fluorescence of solid solutions MgO-NiO (Physica **18**, 77-82, 1952, No. 2).

The absorption spectrum of MgO-NiO solid solutions, between 2500 Å and 8000 Å, shows five discrete bands with maxima at 2900 Å, 4000 Å. 4650 Å, 6700 Å and 7200 Å. At liquid air temperatures the 4000 Å and 4700 Å bands show a fine structure with $\Delta\sigma = 235\text{-}240\ \text{cm}^{-1}$. With cathode-rays MgO containing 10^{-3} mole NiO shows a green luminescence at liquid air temperature, in a band consisting of at least eight equidistant sub-bands with $\Delta\sigma = 194\ \text{cm}^{-1}$.

The bands are attributed to electronic transitions between various states of the Ni^{2+} ion, broadened by coupling with lattice vibrations. This explains why NiO exposed to radiations of wavelength 2500 Å $<\lambda<$ 8000 Å does not show photoconductivity.

2034: W. J. Oosterkamp and J. Proper: Free-air and thimble ionization chambers for Grenz-ray dosimetry (Acta Radialogica **37**, 33-43, 1952, No. 1).

A free-air ionization chamber is described for roentgen radiation with a half value length (H.V.L.) between 0.02 and 1 mm Al. A dose rate of 50 000 r/min can easily be measured.

Rigid thimble ionization chambers can be made with relatively thick windows (0.2 mm and 0.45 mm respectively). Their sensitivity is independent of wavelength between 0.02 and 1 mm Al and between 0.04 and 2.5 mm Al H. V. L. respectively. The independence is obtained by using for the wall a material that has an effective atomic number lower than that of air. This is arranged by a suitable combination of window thickness, depth of the ionization chamber and effective atomic number of the wall material. The variation, with the radiation quality, in the attenuation of the beam due to the chamber window and in the contribution to the ionization from secondary electrons released in the chamber wall, compensate each other.

Dose rate and absorption curves in aluminium, measured with these ionization chambers, are given for a tube with mica-beryllium window (inherent filter equivalent to 40 μ Al) for tube voltages of 11 kV, 16.5 kV, 21.5 kV, 31.0 kV and 49.5 kV (DC).

2035: N. Warmoltz and E. Bouwmeester: An easily degassable ionization gauge with a simple and stable circuit (Appl. sci. Res. **B 2**, 273-276, 1952).

An ionization gauge of the tetrode type is described, with two grids and a thin ion collector on the glass wall. The tube is very easily degassed by heating it in an oven and passing a current through the grids.

The electron current is stabilized by the first grid. For measuring the ion current a balanced bridge D.C. amplifier with one tube is used.